



Application d'algorithmes prédictifs à l'identification de niches écoculturelles des populations du passé : approche ethnoarchéologique

Nicolas Antunes

► To cite this version:

Nicolas Antunes. Application d'algorithmes prédictifs à l'identification de niches écoculturelles des populations du passé : approche ethnoarchéologique. Archéologie et Préhistoire. Université de Bordeaux, 2015. Français. NNT : 2015BORD0173 . tel-01359190

HAL Id: tel-01359190

<https://theses.hal.science/tel-01359190>

Submitted on 2 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

présentée au Laboratoire PACEA pour
obtenir le diplôme de doctorat

Spécialité : **Préhistoire**
École Doctorale : **Sciences et Environnements**

Application d'algorithmes prédictifs à l'identification de niches écoculturelles des populations du passé : approche ethnoarchéologique

par

Nicolas ANTUNES

Soutenue le 6 novembre 2015, devant le jury composé de :

Président du jury

Jacques JAUBERT, professeur Université de Bordeaux, Bordeaux

Directeur de thèse

Francesco D'ERRICO, directeur de recherche CNRS, Bordeaux

Co-Directeur de thèse

William E. BANKS, chargé de recherche CNRS, Bordeaux

Rapporteurs

Michel CRUCIFIX, professeur Université de Louvain, Belgique

A. Townsend PETERSON, professeur Université du Kansas, USA

Examineurs

Christophe COUPÉ, chargé de recherche CNRS, Lyon

Pierrick LEGRAND, maître de conférences Université de Bordeaux, Bordeaux

Wulf SCHIEFENHÖVEL, professeur Institut du Max-Planck, Seewiesen, Allemagne

*Intelligence is sometimes defined
as the capacity of the individual to
adjust himself successfully to his
environment or to adjust the
environment to his needs*

L'intelligence est parfois définie
comme la capacité d'un individu
à s'adapter à son environnement
avec succès ou à ajuster son
environnement à ses besoins.

BRUCE LEE



Remerciements

J'écris ces quelques lignes afin de remercier toutes les personnes qui ont contribué de près ou de loin à l'accomplissement de cette thèse ainsi que tous ceux qui m'ont accompagné scientifiquement et moralement durant ces quatre années.

Avant tout, je tiens à remercier chaleureusement Francesco d'Errico et William Banks pour avoir accepté de codiriger cette thèse. Merci pour votre investissement, vos disponibilités, vos corrections, vos conseils avisés et votre amitié. Merci de m'avoir donné l'opportunité d'exercer dans le milieu de la Préhistoire : ma passion. Merci de m'avoir fait confiance.

Je suis très honoré que le professeur Townsend Perterson ait accepté de rapporter cette thèse. Son avis en tant que spécialiste des modélisations de niches écologiques est tout aussi précieux qu'indispensable et ses remarques pertinentes vont me permettre d'améliorer ce travail.

Je remercie aussi vivement le professeur Michel Crucifix d'avoir également été le rapporteur de ce document, en effet les données climatiques ont une place importante dans ce travail et par conséquent son aval compte beaucoup.

Je suis touché par la présence dans le jury du professeur Wulf Schiefenhövel. Sa passion à transmettre généreusement son savoir et sa remarquable expérience acquise suite à de nombreuses années de terrain passées en Nouvelle-Guinée en font une personne que j'admire particulièrement. Wulf, je vous remercie également pour votre hospitalité tant à Andechs qu'au Tyrol.

J'adresse tous mes remerciements à Pierrick Legrand pour sa participation au jury.

J'exprime toute ma gratitude au professeur Jacques Jaubert pour avoir accepté également de juger ce travail.

Je suis extrêmement reconnaissant à Christophe Coupé pour sa participation à mon jury de thèse. Son expertise en tant que linguiste en fait un excellent interlocuteur.

Ce travail n'aurait pu s'accomplir sans Solange Rigaud et Marian Vanheren avec qui j'ai collaboré et qui, au fil du temps sont devenues de véritables amies. Sans vos participations cette thèse n'aurait certainement pas pu aboutir. Merci pour votre aide, votre dévouement, votre réactivité, votre perspicacité et votre gentillesse. Solange, c'est ta rencontre lors d'un chantier de fouilles qui m'a conduit jusqu'au laboratoire, je t'en suis reconnaissant. Marian, le fait de m'avoir intégré à ton projet de recherche a considérablement enrichi mon travail de thèse ; de plus, travailler avec toi est un vrai plaisir, mille mercis.

Je remercie Guillaume Levavasseur et Marie Noelle Woillez de m'avoir initié au *downscaling statistique*.

Merci Jérôme Servonat et Didier Swingedou d'avoir pris le temps de m'expliquer comment manipuler les données issues de simulations climatiques.

Merci à Richard Michalet de m'avoir consacré un peu de temps afin de discuter de notions d'écologie, ce fut très instructif.

Merci Aurélien Royer, notre travail sur les niches écologiques des lemmings a été très stimulant, nous devons absolument le continuer. Qui aurait cru qu'en croquant dans un sandwich on puisse avoir une telle idée ! Mais surtout merci Chapô pour ton amitié.

Merci Serena Tucci et Andrea Brunelli, vos expertises combinées aux nôtres vont donner naissance à un projet stimulant, je suis impatient de continuer ce travail que nous avons entamé.

Je remercie mes proches qui m'ont soutenu et aidé particulièrement à la fin, lorsqu'il fallait coucher sur le papier mes travaux. Un remerciement particulier à Younès Darradi qui m'a apporté une aide considérable et dont je salue la patience. Je tiens ici à exprimer ma reconnaissance envers d'autres amis qui m'ont également relu : Séverine Fabre, Johanna Karsenty, Émilie Campmas, Adrien Labone, Solange Rigaud, Philippe Juan, Olivier Legardeur, Marian Vanhaeren, Léonie Rey, Didier Swingedouw, Angélique Villa, Adeline Launey (grenouille), Jean-Renaud Despres. Merci Floréal Morandat pour ton aide et pour m'avoir enseigné comment se servir de ta classe \LaTeX que j'ai utilisée pour rédiger ce document.

L'archéologie c'est avant tout du terrain. La promiscuité lors d'un chantier avec des spécialistes en divers branches de la discipline permet un éveil scientifique hors du commun. Je remercie particulièrement Marie Soressi, de m'avoir offert l'opportunité de fouiller pour la première fois et ainsi me donner la possibilité de confirmer mon goût pour la Préhistoire. Je remercie également les autres directeurs de fouilles qui m'ont accueilli ultérieurement sur leurs chantiers : Cédric Beauval, William Rendu et Vincent Mourre. Je profite ici pour saluer tous les fouilleurs que j'ai pu croiser sur ces divers chantiers.

Cette thèse a été réalisée dans les locaux du laboratoire PACEA, précisément dans la partie où sont réunis les Préhistoriens. Je remercie Bruno Maureille, directeur de l'unité, de m'avoir accueilli dans ce laboratoire. Les années passées dans ces locaux ont été agréables. Merci aux doctorants de l'époque, qui ont maintenant soutenu, de m'avoir si bien accueilli dès mon arrivée en master : Emmanuel Discamp, Luca Sitzia, Myriam Boudadi-Maligne, Solange Rigaud, Katja Douze, Émilie Campmas. En ce sens j'exprime aussi ma gratitude envers les autres membres de l'Unité pour leur convivialité : Michèle Charuel, Eric Pubert, Émilie Claud, Catherine Morel-Chevillet, Françoise Lagarde, Isabelle Esqurial, Dominique Armand, Anne Delagnes, Jean-Luc Guadelli, Aleta Guadelli, Michel Lenoir, Alain Queffelec, Catherine Ferrier, Julia Roussot-Larroque, Mathieu Langlais, Véronique Laroulandie, Jean-Baptiste Mallye, Arnaud Lenoble, Patrick Michel, Pascal Berthran, Jean-Guillaume Bordes, Jean-Philippe Faivre, Sylvain Ducasse.

Un merci particulier aux doctorants avec qui j'ai partagé le bureau au fond de cet interminable couloir et ceux du bureau voisin qui ont su me dépanner, me conseiller mais aussi me supporter. Merci Tiphaine Maurin, Dounia Large, Daniela Rosso, Solène Caux, Renaud Lippé, Eric Andrieux, Marie-Claire Dawson, Brad Gravina, Anna Madji, Anthony Secher, Michael Baillet, Alexandre Lefebvre. Nos moments de détente lors des pauses café et déjeuner et nos sorties en soirées m'ont apporté beaucoup. Je remercie également par la même occasion les post-doctorants et les étudiants de master (pour la plus part en thèse aujourd'hui) qui ont également partagé ces moments conviviaux : Africa Pitarch Marti, Elysandre Puech, Léa Feyfant, Lysianna Ledoux, Sofia Solanas, Elsa Garot, Marie Matu, Maxime Pelletier, Caroline Masset et James Blinkhorn (Jimbob).

Ces quelques lignes sont aussi l'occasion de remercier les personnes que j'ai côtoyées pendant ces années de thèses et qui m'ont apporté un soutien considérable. Je conclurai donc en remerciant tous mes amis proches qui représentent à mes yeux ma famille.

- Merci ma pote Emilie, sacrée déconneuse, héhéhé.
- Merci ma pote Katja, *be water my friend*.
- Merci mon pote Younès, l'extraordinaire maroco-charentais.
- Merci mon pote Alberto, que dirait Napoléon dans cette situation ?
- Merci la bande bordelaise : Vincent F., Marion, Flop, Maf, Ralu, Philou, Élisabeth, Em' et Em', Soizic, Marius, Dennis, Jean-Paul, Vincent M., Sophie, Sandra.
- Merci la bande toulousaine : Jérôme, David, Johanna, Romain, Élo, Angel, Dimitri, Marlène, Sève, Yann, Yack, Gaëlle, Aline - tu nous manques.
- Merci la bande bornéenne : Olive, Élodie, Mich, Renaud, Mars, Seb.

Enfin, merci Man' de m'avoir relu et de m'avoir toujours soutenu dans mes choix.

Cette recherche a pu être menée grâce à une bourse ministérielle et aux financements des ERC TracSymbols et du projet LascArBx Ethnoparure.



Table des matières

Remerciements	iii
Table des matières	vii
Guide du document	1
Termes scientifiques	1
Glossaire	2
Introduction	3
Présentation du manuscrit	5
Structure de la thèse	5
I Modélisation de niches écologiques : concepts et méthodologie	7
1 Théories et concepts écologiques	9
1.1 Cadre théorique des modélisations de distribution de populations	9
1.2 Définitions et nuances entre distribution, habitat et niche écologique	11
1.2.1 Distribution géographique d'une espèce	11
1.2.2 Habitat d'une espèce	14
1.2.3 Niche écologique d'une espèce	15
1.2.4 Le diagramme BAM	17
1.3 Dynamiques des distributions d'espèces.	18
1.3.1 Interactions abiotiques	19
1.3.2 Interactions biotiques	19

1.3.3	Vision moderne de la niche écologique	24
1.3.4	Niche ou habitat deux ensembles non comparables.	24
1.4	Modélisation de niche ou d'habitat, que modélise-t-on?	25
1.4.1	Impact des interactions biotiques sous-jacentes dans l'élaboration des modèles	26
1.4.2	L'espace écologique modélisé	28
1.5	Évolution constante des niches écologiques	33
1.6	Concept de niches écoculturelles	34
1.6.1	Définition de la niche écoculturelle	34
1.6.2	Principe de l'ECNM	35
1.6.3	Répartition géographique des groupes humains	35
1.6.4	Substrats environnementaux, expressions culturelles et cohésion so- ciale	36
1.6.5	Un rythme d'évolution rapide	37
1.6.6	Portées et limites de l'ECNM	38
1.7	Exemples d'applications d'algorithmes prédictifs	41
1.7.1	Utilisation des algorithmes prédictifs en écologie et paléoécologie . .	41
1.7.2	Utilisation des algorithmes prédictifs pour des groupes humains . . .	42
2	Les algorithmes prédictifs de niches écologiques	47
2.1	Modélisation de niches écologiques à l'aide d'algorithmes prédictifs	48
2.2	Les Algorithmes prédictifs en détails	48
2.2.1	Fonctionnement général des algorithmes prédictifs	48
2.2.2	Prédictions et projections de niches	52
2.2.3	Liste non exhaustive des algorithmes prédictifs, descriptions et spéci- ficités	52
2.2.4	Caractère stochastique des prédictions	56
2.2.5	Les algorithmes préférentiellement utilisés	56
2.2.6	Les algorithmes les plus adaptés aux études archéologiques	59
2.3	Représentativité des données	59
2.3.1	Qualité des points d'occurrences	59
2.3.2	Qualité des données topographiques	63
2.3.3	Qualité des données climatiques	65
2.4	Traitement des données avant la modélisation	68
2.4.1	Traitement des occurrences	68
2.4.2	Traitement des variables environnementales	68
2.5	Optimiser une prédiction	70
2.5.1	Utiliser plusieurs jeux de variables environnementales	71
2.5.2	Les modélisations seuillées	71
2.5.3	Quels seuils appliquer en ECNM?	72
2.6	Stochasticité, validité et pertinence des prédictions	73

2.6.1	La matrice de confusion	73
2.6.2	Évaluer une prédiction	74
2.6.3	Tests de performance des modèles	76
II	Développement d'outils spécifiques à la modélisation de niches	83
3	Collecter les occurrences	85
3.1	Méthodes de collectes classiques	85
3.2	Affiliation à partir de régions prédéfinies	86
3.3	Extractions d'occurrences à partir de cartes scannées	87
4	Obtenir des couches environnementales haute définition pour une période spécifique	89
4.1	Modèles climatiques globaux	89
4.2	Modèles climatiques régionaux	90
4.3	<i>Downscaling</i>	90
4.4	Descente d'échelle statistique par modèle additif généralisé	92
4.4.1	La modélisation GAM	92
4.5	Critiques de la méthode présentée	104
5	Automatisation de la modélisation de niches	105
5.1	Intérêt de multiplier les modèles	105
5.2	État de l'art	106
5.3	Scripts de génération de commande <i>om_console</i>	106
6	Création de consensus de modèles de niches	109
6.1	Multiplier les techniques de modélisation	109
6.2	Présenter les résultats issus de plusieurs techniques de modélisation	110
6.2.1	Sélection d'algorithmes	110
6.2.2	Prédictions consensuelles	112
6.2.3	Comparaison consensus classiques et consensus élitistes pondérés	115
6.3	Sélection ou consensus : que choisir ?	116
6.4	Conception de consensus élitistes pondérés	117
6.4.1	Rendre comparables les prédictions	117
6.4.2	Exemple de création d'un consensus élitiste pondéré	118
7	Analyse des niches écologiques	123
7.1	Regard critique sur les données	123
7.2	Examiner la répartition géographique des prédictions	124
7.2.1	Comparaison avec nos connaissances sur la culture	124
7.2.2	Comparaison avec l'environnement	124

7.3	Mesures et comparaison de niches	124
7.4	Identification et quantification des principaux facteurs environnementaux	125
7.4.1	Les « ACP communes »	126

III Reconstructions de niches écoculturelles, trois cas d'études 129

8	Impact de l'environnement sur la diversité linguistique de Nouvelle-Guinée	131
8.1	Relations langue/ethnité/environnement	132
8.1.1	Diversité linguistique/ethnité	132
8.1.2	Diversité linguistique/environnement	132
8.2	Contexte environnemental de Nouvelle-Guinée	135
8.2.1	Géographie physique et climat	135
8.2.2	Diversité environnementale de Nouvelle-Guinée	137
8.3	Histoire du peuplement de la Nouvelle-Guinée, environnement et diversité humaine	139
8.3.1	Le peuplement papou	140
8.3.2	Le peuplement Austronésien	143
8.3.3	Impact de l'environnement biotique sur les géotypes humains et le peuplement	145
8.4	Contexte linguistique de Nouvelle-Guinée	146
8.4.1	Une diversité linguistique très élevée	146
8.4.2	Phylolinguistique de Nouvelle-Guinée	146
8.5	Modélisation des niches écolinguistiques de Nouvelle-Guinée	148
8.5.1	Occurrences	148
8.5.2	Données environnementales	150
8.5.3	Délimitation du M	150
8.5.4	Seuil d'omission	151
8.5.5	Consensus élitistes pondérés	151
8.5.6	Résultats de l'ELNM	151
8.5.7	Différences Papous-Austronésiens	179
8.6	Perspectives	183
9	Impact du changement climatique sur les populations scandinaves du Groenland	187
9.1	La détérioration climatique du PAG	188
9.2	L'histoire des Vikings des régions subarctiques	189
9.2.1	Les Vikings, une épopée nordique	189
9.2.2	Redoutables guerriers, mythe ou réalité ?	189
9.2.3	Témoignages écrits	190
9.2.4	Vestiges archéologiques	191
9.2.5	Le <i>Landnám</i> ou la colonisation des terres nordiques	191

9.2.6	L'économie nordique	194
9.2.7	Déclin d'une civilisation	195
9.3	ECNM des Vikings des régions subarctiques	196
9.3.1	Données archéologiques	196
9.3.2	Étude de faisabilité	197
9.3.3	Une ECNM plus élaborée	200
9.4	Interactions culturelles	223
10	Modélisation des niches écoculturelles des cultures initiales du Néolithique	225
10.1	Les grands courants de néolithisation en Europe	225
10.2	Niches écoculturelles des cultures du Néolithique ancien	228
10.2.1	Matériels et méthodes	228
10.2.2	Principaux résultats et discussion	231
10.3	Nouvelle étude	232
10.3.1	Matériels et méthodes de la nouvelle étude	233
10.3.2	Résultats de la nouvelle étude	234
10.4	Perspectives apportées par la nouvelle étude	243
	Conclusions et perspectives	247
	Synthèse générale	247
	Optimisation de l'ECNM	247
	Démarche ethnoarchéologique	248
	Apports de l'ECNM	249
	Discussion et perspectives	249
	Diversité culturelle	249
	Évolution, compétition et peuplement	250
	Ethnicité des populations archéologiques	251
	Bibliographie	253
	Table des figures	297
	Liste des tableaux	301



Guide du document

Termes scientifiques

Des termes anglais ont été employés dans cette thèse, ils se trouvent alors écrits en italique et suivent les règles grammaticales de cette langue.

Termes statistiques. Certains mots anglais pour désigner des méthodes statistiques ont été utilisés bien que leur équivalent français existent ou qu'ils puissent être traduits. Dans un souci de clarté et d'homogénéité avec la littérature scientifique nous avons préféré garder la dénomination la plus courante. Ainsi nous garderons les formes françaises pour les opérations les plus basiques et utiliserons la dénomination anglaise pour les plus spécifiques à la modélisation de niches écologiques.

Toponymes et noms propres. Dans le chapitre 8, beaucoup de groupes linguistiques présentés n'ont pas de documentation française. Nous avons donc été contraints de conserver le toponyme anglais qui est utilisé dans la littérature scientifique sur laquelle nous nous appuyons. Afin de respecter une homogénéité, nous avons donc décidé d'employer les toponymes anglais pour l'ensemble des groupes de langues que nous traitons dans ce chapitre. Les noms propres et les adjectifs anglais faisant référence à une langue, une région ou un groupe de langues commenceront par une majuscule et n'auront pas de forme plurielle.

Sigles et acronymes anglais. Dans la majorité des cas, nous préférons utiliser des sigles et acronymes anglais mais nous plieront ces derniers aux règles de la grammaire française (pas de marque du pluriel) et ne les écrirons pas en italique.

Termes spécifiques à la thèse. Dans une optique de cohérence avec nos futurs travaux et publications, nous estampillons les méthodes développées en anglais.

Typographie. Les données chiffrées présentées dans les tableaux et graphiques ont un point comme séparateur de décimale. Dans un souci d'homogénéité, les séparateurs de décimales du corps de texte seront également des points.

Glossaire

AD :	<i>Anno Domini</i> , système de datation, signifiant après Jésus Christ.
AUC :	<i>Area Under the Curve</i> , niche écoculturelle.
BP :	<i>Before Present</i> , système de datation, signifiant avant le présent (1950).
DMG :	Dernier Maximum Glaciaire
ECN :	<i>Ecocultural Niche</i> , niche écoculturelle.
ECNM :	<i>Ecocultural Niche Modeling</i> , modélisation de niches écologiques.
ELN :	<i>Ecolinguistic Niche</i> niche écolinguistique.
ELNM :	<i>Ecolinguistic Niche Modeling</i> , modélisation de niches écolinguistiques.
ENH :	<i>Eltonian Noise Hypothesis</i> , hypothèse du bruit eltonien
ENM :	<i>Ecological Niche Modeling</i> , modélisation de niches écologiques.
ma :	Millions d'années.
ka :	Milliers d'années.
OCM :	Optimum climatique médiéval.
PAG :	Petit Âge Glaciaire.
PPN :	Production primaire nette.
ROC :	<i>Receiver Operating Characteristic</i>
SDM :	<i>species distribution modeling</i>
TSS :	<i>True Skill Statistic</i>



Introduction

L'ÉVOLUTION des comportements humains face aux changements climatiques est un sujet d'actualité. Les traits de côte se retracent, des écosystèmes entiers sont perturbés engendrant l'extinction de nombreuses espèces animales ou végétales, la banquise du Pôle Nord diminue et la liste est encore longue (rapport de l'IPCC, [Field et Van Aalst, 2014]). Ces changements modifient nos environnements et par conséquent nos systèmes économiques, nos modes de vie. Nos connaissances actuelles sur l'émission de gaz à effet de serre nous encouragent, au moins par précaution, à changer nos habitudes et nos comportements.

Par le passé, les sociétés humaines ont vécu, avec des ampleurs variables, des changements climatiques qui les poussèrent à migrer, s'adapter, innover, se transformer ou bien disparaître [Gronenborn, 2003; McLeman et Smit, 2006; Smithers et Smit, 1997]. L'Histoire et la Préhistoire nous dévoilent peu à peu, par le biais de nouvelles techniques et découvertes, les solutions fructueuses ou non adoptées par les Hommes du passé. Ainsi les travaux des archéologues, qui rassemblent des éléments et recueillent des informations sur la pérennité des cultures du passé, apportent un éclairage quant à nos choix et nos actions du quotidien, lesquels jouent un rôle dans l'avenir de nos traditions et modes de vie actuels.

L'archéologie est une science humaine qui s'attache à identifier, par le biais de vestiges matériels, des comportements sociétaux. Il a été identifié une grande diversité de comportements adoptés par les Homininés depuis la Préhistoire ancienne, et ce bien avant l'apparition d'*Homo sapiens*. En premiers lieux, les comportements se distinguaient par des nuances dans les stratégies de subsistance, dans les structures sociétales et fonctionnelles [Binford, 2001; Speth, 1987]. Par la suite des notions plus spirituelles avec l'émergence du symbolisme et du religieux sont nées [d'Errico et Sánchez Goñi, 2003; Stark et Bainbridge, 1987]. Il en résulte des diversités culturelles unissant ou séparant les groupes humains en leur conférant des sentiments d'ethnicité auxquels ces groupes sont peu ou prou attachés.

Un grand débat existe entre les « naturalistes » qui estiment que la diversité culturelle

est déterminée par l'environnement et les « culturalistes » qui au contraire jugent que l'évolution culturelle est indépendante de l'environnement. Les deux courants de pensées évoquent des arguments robustes laissant présager que la diversité culturelle doit dépendre des deux processus évolutifs.

[Pagel et Mace \[2004\]](#) stipulent qu'en dépit de constituer une espèce relativement homogène génétiquement, les Hommes sont particulièrement aptes à trouver des différences entre eux, les divisant en groupes ethniques. Les différences sociales, linguistiques ou, de manière générale, culturelles peuvent générer ainsi des exclusions sympatriques [[Erikson, 1996](#)]. Ces divergences culturelles entre groupes humains découlent-elles de processus aléatoires ? Sont-elles initiées ou favorisées par des conditions environnementales spécifiques ? Ont-elles lieu avec la même intensité dans tous les milieux ? Les humains ont développé des technologies adaptées leur permettant de s'affranchir des contraintes environnementales qui leur sont imposées. Ainsi des civilisations ont vu le jour sous toutes les latitudes. Chacune possédant ses propres caractéristiques fonctionnelles et culturelles. Chacune adaptée à son environnement. Cependant le climat a subi d'importantes variations depuis l'aube de l'humanité. Quelles ont été les conséquences de ces fluctuations sur les différents modes de vie, les différentes cultures du passé ? Ont-elles toujours eu un impact de même intensité ? Sont-elles causes d'effondrements culturels ou n'ont-elles joué qu'un rôle minime face aux facteurs socio-économiques ? Mesurer la part de l'environnement et du culturel dans l'établissement d'un groupe humain et déterminer les barrières écologiques ou culturelles mises en place apportent des éléments de réponses à ces questions.

L'histoire du peuplement et de l'installation des groupes humains peut être abordée selon différentes approches. Les recherches actuelles s'intéressent tant à la diversité biologique et linguistique des populations qu'à leur culture matérielle. Un des grands enjeux est de contextualiser ces données dans leur cadre environnemental. La modélisation de niches écoculturelles [[Banks *et al.*, 2006](#)] consiste en l'utilisation d'algorithmes prédictifs de niches écologiques pour croiser des données environnementales et culturelles. Ces dernières années, plusieurs études ont montré l'intérêt d'appliquer cette méthode au registre fossile (par exemple [[Banks *et al.*, 2011, 2008b,c, 2013b, 2009](#)]) afin d'apprécier la distribution potentielle de populations humaines possédant le même système technique ou les mêmes traits culturels. Ces travaux ont montré que l'utilisation d'algorithmes prédictifs permet l'identification de liens cultures/environnements. Cette approche permet aussi la quantification des paramètres les plus influents et la mise en évidence des phénomènes de décalage, de contraction ou d'expansion de niche lorsque ces dernières sont projetées dans différentes phases climatiques [[d'Errico et Banks, 2013](#)]. Cependant la modélisation de niche écoculturelles des périodes passées se heurte à deux difficultés majeures. La première se situe au niveau de la précision, de la disponibilité, de la qualité et de la quantité des données. La seconde tient du fait qu'il s'agit de prédictions dont on ne possède pas la possibilité de vérifier l'exactitude puisque les sujets de la modélisation n'existent plus.

Dans la continuité de ces approches, les travaux présentés dans ce document proposent une optimisation de la méthodologie qui consiste d'une part à augmenter la qualité des

données et à améliorer les performances de la modélisation ; et d'autre part à identifier des dynamiques culturelles de deux populations témoins, l'une actuelle et l'autre historique. Nous verrons ensuite si les schémas comportementaux mis en évidence par l'étude de ces populations peuvent être décelés par l'analyse des niches écoculturelles de populations dont nous ne connaissons pas les relations sociales.

Présentation du manuscrit

Le sujet de cette thèse consiste en l'utilisation d'algorithmes prédictifs afin d'identifier et de quantifier d'éventuelles relations entre cultures et environnements. Le document est divisé en trois parties.

Structure de la thèse

La première partie du document expose les différentes théories écologiques et introduit le concept de niche ainsi que la modélisation de niches écologiques et écoculturelles dans un premier temps, puis décrit dans un second temps le fonctionnement des algorithmes prédictifs de niches.

La deuxième partie présente les outils méthodologiques développés dans le cadre de la thèse. Sont détaillés successivement les méthodes pour :

- collecter des données d'occurrences ;
- obtenir des variables environnementales en accord avec les problématiques de modélisation de niches ;
- automatiser la modélisation de niche ;
- réaliser des consensus de prédictions de niches ;
- analyser les résultats issus de la modélisation.

La troisième partie est composée de cas d'études volontairement présentés dans l'ordre antéchronologique. En effet, comme toute démarche ethnoarchéologique, nos hypothèses se fondent sur des données robustes, précises et vérifiables issues de l'ethnologie afin de les appliquer ensuite, par actualisme, à des périodes plus anciennes. Trois cas d'études sont présentés.

- Le premier cas d'étude porte sur la relation « diversité linguistique/environnement » en Nouvelle-Guinée, il constitue un référentiel ethnographique.
- Le deuxième cas d'étude traite de la disparition des colonies vikings du Groenland et constitue un référentiel historique.

- Le troisième cas d'étude est l'application au Néolithique de la méthodologie présentée dans cette thèse et validée par les deux référentiels.

Enfin, nous discuterons des nouvelles perspectives rendues possibles à partir des éléments présentés dans cette thèse pour l'archéologie de la Préhistoire.

PREMIÈRE PARTIE

MODÉLISATION DE NICHES ÉCOLOGIQUES : CONCEPTS ET MÉTHODOLOGIE

Théories et concepts écologiques

Préambule

Ce chapitre introduit et définit les termes et concepts écologiques qui sont mentionnés tout au long du document. Il décrit les différents types d'interactions entre espèces pour l'accès aux ressources et propose une lecture des niches modélisées en fonction des biais implicites dus aux données initiales.

Sommaire

1.1	Cadre théorique des modélisations de distribution de populations	9
1.2	Définitions et nuances entre distribution, habitat et niche écologique . . .	11
1.3	Dynamiques des distributions d'espèces.	18
1.4	Modélisation de niche ou d'habitat, que modélise-t-on ?	25
1.5	Évolution constante des niches écologiques	33
1.6	Concept de niches écoculturelles	34
1.7	Exemples d'applications d'algorithmes prédictifs	41

1.1 Cadre théorique des modélisations de distribution de populations

Il existe de nombreuses façons de définir une espèce, mais pour les descriptions à venir, nous avons choisi la définition de [Mayr \[1942\]](#), où une espèce est définie comme un ensemble d'individus pouvant se reproduire entre eux et dont la descendance reste féconde, dans des conditions naturelles, sur de nombreuses générations. Afin de présenter

le plus clairement possible les différents concepts et mécanismes écologiques abordés dans cette thèse, les exemples se placent au niveau de l'espèce pour une description théorique, gardant à l'esprit que bon nombre de ces théories reste vrai que l'on se place à un autre rang taxinomique ou que l'on procède à des regroupements d'une autre nature ¹.

La biodiversité représente l'assemblage d'espèces (ou communauté) vivant en un lieu donné. Elle est mesurée en termes de richesse spécifique (nombre d'espèces présentes en ce lieu) et d'abondance d'individus parmi chaque espèce. La biodiversité varie selon les environnements, pour expliquer les variations de présence ou d'absence d'espèces des différents milieux plusieurs théories s'affrontent. Certaines, comme la théorie des niches, expliquent la répartition des espèces par des facteurs écologiques tandis que d'autres, comme la théorie neutraliste, reposent exclusivement sur des principes de dynamiques de populations non liées à l'environnement.

- **Théorie des niches :** elle part de l'idée que les espèces animales ou végétales occupent préférentiellement des lieux en accord avec leur écologie. Il s'agit d'un constat très ancien [Pocheville, 2010]. Cette idée est ensuite formulée plus précisément au début du xx^e siècle [Elton, 1927; Grinnell, 1904, 1917]. La théorie des niches repose sur les concepts de niche écologique et d'habitat ² définie en détail par Hutchinson [1957], Chase et Leibold [2003], Whittaker *et al.* [1973] ou encore Hall *et al.* [1997].
- **Théorie sources-puits :** [Holt, 1985; Pulliam, 1988] Des zones nommées « sources », où règnent des conditions optimales (ou du moins très favorables) à la réalisation de la niche de l'espèce, permettent la présence d'individus dans des zones limitrophes hors de leur niche appelées « puits ».
- **Théorie neutre de l'évolution :** cette théorie purement mécaniste fut énoncée en 1968 par Kimura *et al.* et transférée à l'écologie par Hubbell [2006] (dans [Pocheville, 2010]). Hubbell attribue les nuances de biodiversité aux phénomènes migratoires et à l'apparition de mutations. Son hypothèse, inspirée par la dérive et les modèles neutres rencontrés en génétique, part du principe que toutes les espèces possèdent la même *fitness* ³ et que par conséquent leur distribution n'est due qu'à des phénomènes stochastiques liés aux dispersions.
- **Théorie de la biogéographie insulaire :** [MacArthur, 1967; MacArthur et Wilson, 1963] cette théorie mécaniste, fortement dépendante de la démographie, prend néanmoins en considération l'influence de l'environnement sur la *fitness*. Elle oppose des lieux

1. Lorsque nous aborderons les niches écoculturelles (section 1.6) nous verrons qu'il est possible de regrouper des populations humaines partageant un même trait culturel pour en faire une entité remplaçant l'espèce.

2. L'habitat (voir section 1.2), le lieu de vie de l'espèce, est une notion très ancienne confrontée à un biais épistémologique qui lui confère une définition changeante créant ainsi des confusions notamment lorsqu'il est comparé avec la niche écologique.

3. Terme anglais qui comprend les notions de valeur sélective, réussite, chance de survie et de performance de l'espèce.

appropriés à la survie d'un assemblage d'espèces appelés îles ou continents selon leur superficie au reste de l'environnement hostile pour l'espèce, nommés mers ou océans selon la même allégorie. Plus la taille d'une île est importante et plus elle est proche d'un continent, plus elle est susceptible d'être visitée par de nouvelles espèces et voit alors sa biodiversité augmenter.

- **Métapopulations :** [Levins, 1969] « populations de populations », le milieu de vie de l'espèce est morcelé en zones appelées *patches*⁴ voisins. La structure en métapopulation permet des migrations de populations de *patch* en *patch* qui peuvent assurer, dans une certaine mesure, un maintien de la population en cas de perturbation partielle du milieu.

La densité de l'espèce joue un rôle important sur sa distribution. Or, pour mener le travail que nous présentons ici, nous ne possédons pas cette information. De manière à s'affranchir de ce biais, notre raisonnement est donc centré sur les théories permettant de reconstituer et d'expliquer les distributions géographiques indépendamment des densités des espèces. La démarche des modélisateurs de niches écologiques s'appuie sur la théorie des niches pour définir les facteurs environnementaux expliquant la distribution des espèces, tout en gardant à l'esprit les biais induits par la structure ou les caractéristiques des populations étudiées (par exemple : population isolée, métapopulation), la démographie, ainsi que les variations locales imperceptibles par les méthodes qu'ils emploient.

Selon l'hypothèse de la théorie des niches, trois caractéristiques fondamentales de l'espèce permettent de comprendre ses relations avec son environnement et d'expliquer sa biogéographie. Ces caractéristiques sont : la distribution géographique, l'habitat et la niche écologique.

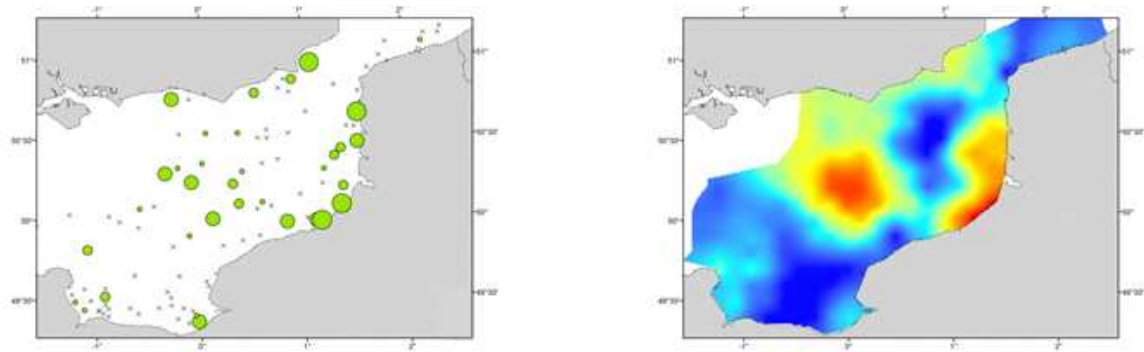
1.2 Définitions et nuances entre distribution, habitat et niche écologique

1.2.1 Distribution géographique d'une espèce

On appelle distribution géographique d'une espèce la région qui englobe toutes les localités où l'organisme a été recensé. Également nommée aire de répartition, la distribution géographique correspond donc à l'étendue spatiale des occurrences d'une espèce donnée. Dans certains cas, cette donnée peut être couplée avec la densité. L'aire représente un ensemble de régions où il est probable de rencontrer l'espèce. On peut alors décrire cette étendue géographique verbalement en dressant, par exemple une liste de localités, ou encore visuellement en traçant sur une carte un polygone dont les sommets sont les occurrences de l'espèce étudiée les plus éloignées. D'autres approches consistent à évaluer

4. Terme anglais, pluriel de *patch*, dont la probabilité d'être occupée est affectée par la probabilité d'occupation des *patches*, pouvant être traduit par parcelle.

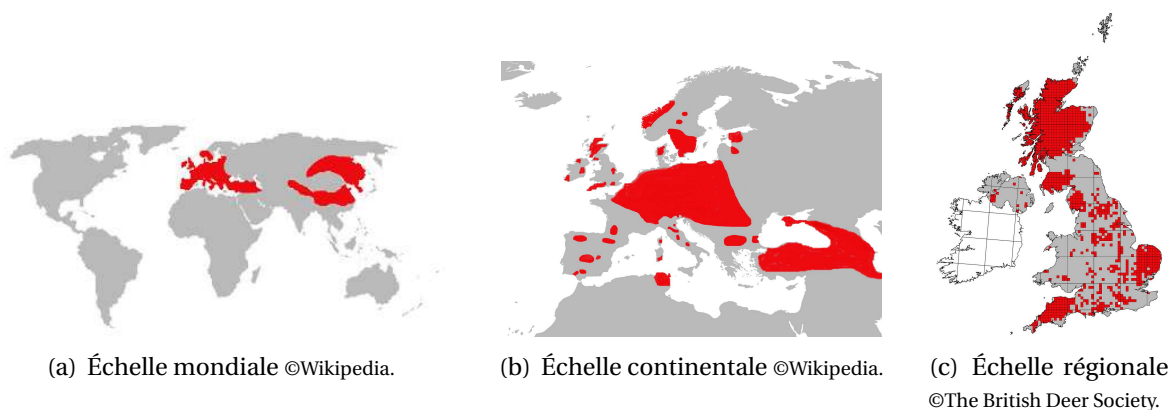
la répartition selon des relevés de fréquences de présence ou à modéliser la distribution de l'espèce. Ce deuxième cas est alors nommé SDM (*Species distribution modeling*). La figure 1.1 présente une densité théorique estimée par interpolation de fréquences de présence alors que la figure 1.2 expose des distributions établies en fonction des occurrences (relevées ou estimées) uniquement. On constate dans ces exemples que les représenta-



(a) Occurrences, la taille des cercles est proportionnelle à la densité.

(b) Interpolation à partir des valeurs de densités ; bleu : densité faible, rouge : densité élevée.

FIGURE 1.1 : Estimation de la densité d'une espèce par Krigeage. L'auteur a réalisé une interpolation à partir d'un comptage d'individus d'une espèce en plusieurs lieux (a). La carte obtenue (b) représente la densité de cette espèce en tous lieux. ©Sandrine Vaz, IFREMER (modifié).



(a) Échelle mondiale ©Wikipedia.

(b) Échelle continentale ©Wikipedia.

(c) Échelle régionale
©The British Deer Society.

FIGURE 1.2 : Distribution géographique de *Cervus elaphus* à différentes échelles, en rouge l'espèce est signalée présente, en gris absente.

tions de distributions admettent différents degrés de finesse selon la précision des relevés d'occurrences. En effet la présence de l'espèce peut être signalée en un lieu très précis

(coordonnées géographiques) ou bien être indiquée selon un découpage administratif : une commune, un pays voire un continent. La figure 1.2 montre la répartition du cerf élaphe (*Cervus elaphus*) à différentes échelles. On s'aperçoit que la précision augmente à mesure que l'échelle s'affine, par exemple, l'Espagne est entièrement comprise dans la distribution de l'espèce représentée au niveau mondial alors qu'à l'échelle européenne, la présence du cerf n'est définie qu'en quelques zones de ce pays seulement. On remarque également l'effet du découpage administratif quand on observe que la présence du cerf n'est pas relevée en Irlande dans la troisième carte alors qu'elle était signalée aux échelles plus grossières.

L'aire géographique d'une espèce est généralement estimée selon l'éloignement géographique des occurrences de l'espèce sans tenir compte de son écologie. Elle inclut des zones où l'espèce ne peut pas vivre, car inaccessibles ou trop pauvres en ressources. Par ailleurs, la distribution géographique d'une espèce à un instant donné dépend fortement de sa mobilité et de son histoire. Ainsi la distribution de l'espèce peut être le résultat d'interactions avec d'autres êtres vivants : présence de compétiteurs, de commensaux, de parasites, introduction de prédateurs, contractions de maladies [Peterson et Cohoon, 1999; Wiens et Graham, 2005] ou de l'impact de facteurs abiotiques : changements climatiques ou émergence de barrières géographiques provoquant une insularisation⁵[Wiens, 2004]. En ce sens, la distribution géographique d'une espèce évolue dans le temps et est le résultat de processus historiques difficilement discernables par de simples études spatiales.

En paléontologie et en archéologie, le registre fossile permet difficilement de reconstituer avec certitude les distributions géographiques et d'en apprécier les dynamiques spatio-temporelles. En effet, le caractère sporadique des données n'est pas adapté pour reconstruire une distribution précise, comme l'illustre la figure 1.3, où l'auteur, disposant de peu d'éléments probants, se trouve dans l'obligation d'émettre plusieurs hypothèses expliquant la diffusion des Homininés. La figure représente les distributions continentales et temporelles de populations basées sur des découvertes en paléanthropologie. En dressant cette figure l'auteur a voulu décrire les principaux *scenarii* évolutifs de la lignée du genre *Homo* les plus probables selon les informations qu'il possédait. Il laisse en suspens plusieurs hypothèses et les résolutions temporelles et spatiales sont très peu précises du fait de la rareté des occurrences.

Sans tenir compte des exigences écologiques de l'espèce, la distribution géographique est une information de qualité variable. Elle peut être très précise lorsque les données sont nombreuses et issues d'un échantillonnage représentatif ou au contraire vague, lorsqu'issue de recensements plus flous (exemple figure 1.3 où la précision géographique est de l'ordre du continent). Cependant la distribution d'une espèce devient une information de qualité quand elle est accompagnée de données environnementales. Le lien entre l'environnement et la distribution de l'espèce nous renseigne sur l'habitat de cette dernière et permet de définir son écologie. Dès lors on ne parle plus de SDM mais de ENM (*ENM* : *Ecological Niche Modeling*) [Peterson et al., 2002]. L'ENM consiste à estimer la biogéographie d'une espèce en

5. Lorsqu'un milieu propice à une communauté est subitement entouré par un milieu qui lui est hostile.

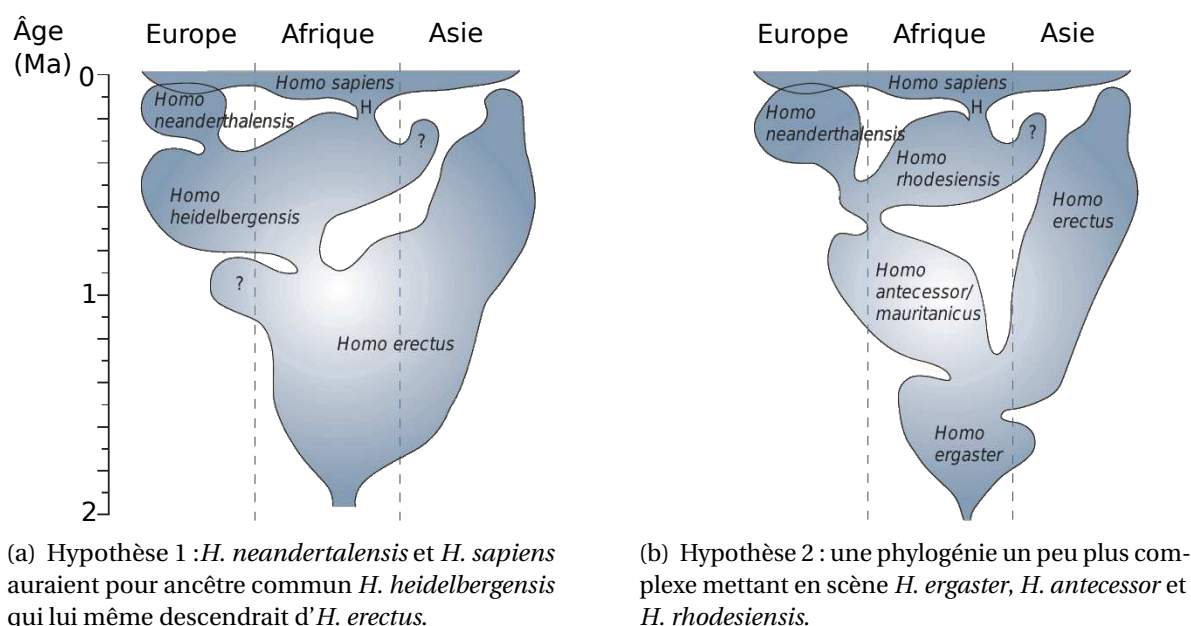


FIGURE 1.3 : Deux hypothèses de distributions spatio-temporelles d'espèces du genre *Homo* basées sur des occurrences paléanthropologiques, d'après [Stringer, 2003] (modifié).

fonction de ses occurrences et des paramètres environnementaux grâce à l'utilisation d'algorithmes prédictifs. Le chapitre 2 est entièrement dédié à l'utilisation de ces algorithmes en macro-écologie.

1.2.2 Habitat d'une espèce

La définition de l'habitat est sujette à polémique. Calenge [2005] constate la rareté de définitions claires [Karr, 1981; Morrison, 2001] et souligne le fait que cette notion, souvent mentionnée mais non, ou vaguement, définie, est généralement considérée comme connue du lecteur [Hall *et al.*, 1997].

L'habitat est l'espace écologique qui définit le milieu où l'espèce vit ou pourrait vivre [Morrison, 2001]. Il est déterminé par son environnement spatial, biotique et abiotique. Il inclut les ressources alimentaires, les abris, ainsi que toutes les conditions environnementales nécessaires à la survie et à la reproduction de l'espèce. Ces conditions étant susceptibles d'évoluer, l'habitat établi à partir de relevés pour une espèce est dépendant de la date et du lieu relatif à l'échantillon qui a été observé pour le définir.

Comme la distribution géographique de l'espèce, l'habitat peut être défini de manière plus ou moins fine selon l'échelle et la précision de la description. À large échelle l'habitat est souvent défini par des biomes ou des paysages spécifiques : région désertique, plaine steppique, savane, etc. [Karr, 1981]. Mais l'habitat peut être décrit plus finement : étage

particulier de la canopée, profondeur de lac d'une région donnée, espèce particulière d'arbres morts à une saison donnée. L'habitat est déterminé à partir d'observations ou de connaissances de l'écologie de l'espèce. Il peut évoluer dans le temps selon les cycles de vie (larves, adultes) ou selon les activités quotidiennes (chasse, nichée). La taille des habitats est variable selon les espèces et ils s'imbriquent souvent les uns dans les autres, certaines espèces constituant des habitats pour d'autres.

1.2.3 Niche écologique d'une espèce

L'habitat et la niche écologique décrivent différemment l'écologie d'une espèce. L'habitat est relatif au milieu et la niche à la position et à la fonction de l'espèce dans l'écosystème. Comme pour l'habitat, la définition de niche écologique présente également des ambiguïtés, dues à des raisons épistémologiques.

1.2.3.1 Niches grinellienne et eltonienne

[Grinnell \[1917\]](#) fut le premier auteur à employer le terme de niche. La niche écologique correspond alors à un milieu pour lequel l'ensemble des paramètres environnementaux permettent la survie de l'espèce. Ces paramètres peuvent être d'ordre abiotique (comme la température, les précipitations, ou l'altitude) ou biotique (présence d'espèces compétitrices, prédatrices, commensales ou facilitatrices). La représentation géographique de la niche grinellienne est la somme de tous les lieux qui présentent les conditions environnementales favorables à l'espèce. Cette niche permet de positionner une espèce dans son milieu de vie par des facteurs environnementaux.

[Whittaker *et al.* \[1973\]](#) opposent la niche grinellienne à la niche eltonienne, car la deuxième renvoie au rôle joué par l'espèce dans l'environnement et à la place qu'elle occupe dans le réseau trophique. Les niches grinellienne et eltonienne positionnent en lieu et fonction l'espèce dans son environnement biotique et abiotique. Ainsi, en se rapportant aux travaux de [Whittaker *et al.* \[1973\]](#), et d'[Hutchinson \[1957\]](#) on peut distinguer deux types de niches grinelliennes : la niche fondamentale et la niche réalisée.

1.2.3.2 Niche fondamentale, niche réalisée et hypervolume d'Hutchinson

La niche écologique fondamentale est décrite par l'ensemble des conditions environnementales abiotiques nécessaires à la survie de l'espèce.

La niche écologique réalisée est définie par le même ensemble de conditions environnementales favorables à l'espèce, mais considère également l'aire qui lui est accessible ainsi que la présence d'autres espèces.

[Hutchinson](#) définit la niche d'une manière qui inspirera et sera le fondement des modélisations de niches écologiques (ENM). L'écologue conceptualise la niche écologique de l'espèce comme un hypervolume où chaque dimension de l'hyperespace est une ressource

de l'environnement affectant sa *fitness*. Tous les paramètres environnementaux en quantité q qui entrent en jeu dans la subsistance d'une espèce à un temps t constituent son écologie. La niche écologique est donc par extension tous les lieux regroupant ces mêmes paramètres en une quantité avoisinant q (en fonction de la tolérance de l'espèce pour les variations de ces paramètres) au même temps t . La *fitness* de l'espèce est maximale au centre de cet hypervolume et diminue à mesure que l'on s'en éloigne. On peut alors tracer son évolution pour chaque dimension environnementale sur un graphique à n dimensions ou chacun des n axes représente une variable environnementale. Souvent, pour simplifier ou théoriser, les réponses de l'espèce aux divers facteurs environnementaux sont rassemblées sur un même axe portant le nom de gradient environnemental. La réponse ou la *fitness* de l'espèce le long de ce gradient est alors représentée par une gaussienne (figure 1.4). Chacune des dimen-

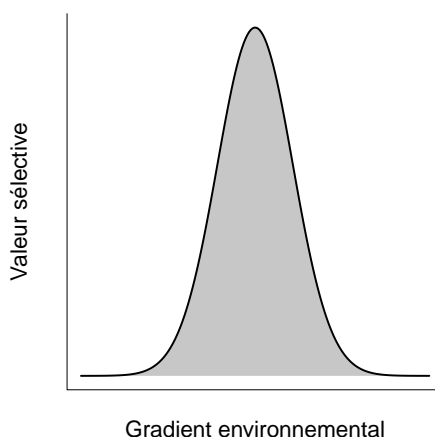


FIGURE 1.4 : Représentation théorique d'une niche écologique. Évolution de la valeur sélective (*fitness*) en fonction du gradient écologique.

sions de la niche écologique peut être analysée séparément (figure 1.5(a)). Il est également possible de regrouper les facteurs sur un même graphique. Une niche à deux dimensions prendra alors la forme d'une ellipse dans un plan constitué des deux facteurs environnementaux (figure 1.5(b)) et une niche à trois dimensions sera représentée par une forme ovoïde (figure 1.5(c)). Suivant ce raisonnement, bien qu'il soit difficile de représenter une niche à plus de trois dimensions écologiques, il est ainsi possible de concevoir mentalement une niche à n dimensions (chacun des n axes représentant un facteur environnemental). Les représentations dans des plans multifactoriels comme ceux des analyses en composante principale permettent également une représentation visuelle des niches.

La conceptualisation de l'hypervolume permet d'entrevoir une analyse de l'influence des facteurs environnementaux un à un comme dans leur ensemble. Cependant il faut noter qu'il est difficile, voire impossible, de dresser une liste exhaustive de ces derniers et donc de modéliser la niche fondamentale exacte. Cette niche reste donc théorique.

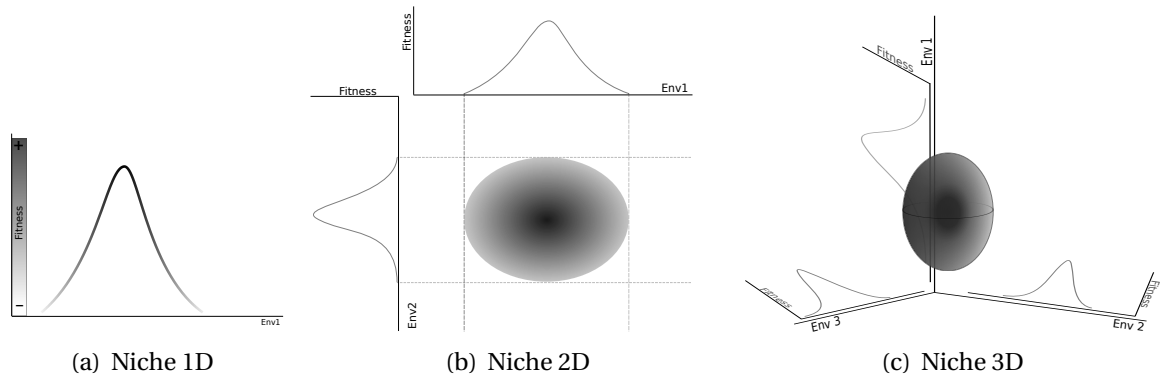


FIGURE 1.5 : Niche écologique 1D, 2D et 3D, évolution de la *fitness* le long de variables environnementales.

1.2.4 Le diagramme BAM

Le diagramme BAM conçoit l'écologie d'une espèce en la replaçant dans un espace selon trois composantes : l'environnement biotique, l'environnement abiotique et l'environnement accessible (figure 1.6). Il s'agit d'un diagramme de Venn qui replace ces trois

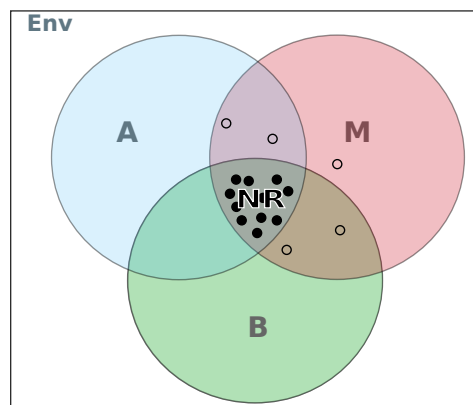


FIGURE 1.6 : Le diagramme BAM, d'après [Soberon et Peterson \[2005\]](#) (modifié). Ce diagramme représente l'enchevêtrement dans l'environnement (Env) des ensembles biotiques (B) et abiotiques (A) dans lesquels peut vivre l'espèce ainsi que l'espace qui lui est accessible reflétant sa mobilité (M). L'intersection $B \cap A \cap M$ est la partie de l'environnement comprenant les conditions strictes requises pour la présence de l'organisme, il s'agit de la niche réalisée (NR). Les cercles représentent les occurrences. Cercles pleins : présences ; cercles vides : absences.

caractéristiques dans l'environnement global. Il permet une réflexion sur la position d'une occurrence parmi ces trois composantes et leurs jonctions. Il offre ainsi la possibilité de clarifier l'espace écologique modélisé.

L'environnement abiotique regroupant les conditions nécessaires à la survie (soit NF , la niche fondamentale de l'espèce) est représenté par l'ensemble A . Les conditions biotiques qui permettent à l'espèce de se développer sont incluses dans l'ensemble B , et les lieux où l'espèce peut accéder sont représentés par l'ensemble M . Hors de chacun des ensembles l'espèce se trouve dans des conditions en désaccord avec son écologie. Les conditions abiotiques impropres à l'espèce sont celles que l'on trouve en dehors de l'ensemble A et forment son espace complémentaire noté \bar{A} . Il en est de même pour la zone inaccessible \bar{M} et l'environnement biotique hostile \bar{B} .

Les occurrences peuvent figurer dans le diagramme BAM (figure 1.6). En théorie l'espèce ne peut se trouver qu'à l'intersection des trois domaines, soit dans sa niche réalisée (NR), car l'organisme ne peut pas vivre ou ne peut accéder au sous ensemble $\bar{B} \cup \bar{A} \cup \bar{M}$.

La NR est une partie de la niche potentielle (NP) définie par certains auteurs comme l'intersection de la niche fondamentale et de l'environnement ($NP_1 = NF \cap E$ soit $NP_1 = NF$) [Jackson et Overpeck, 2000; Soberón et Nakamura, 2009] et par d'autres comme l'intersection des environnements biotiques et abiotiques convenables à l'espèce ($NP_2 = A \cap B = NF \cap B$) [Gaston, 2003; Peterson *et al.*, 2011]. Dans le premier cas (NP_1) l'espèce a accès à tout l'environnement et la condition restrictive à la réalisation de la niche est l'environnement biotique. Dans le second (NP_2) la restriction est d'ordre biotique et géographique. Nous ne raisonnerons par la suite qu'avec la définition de la distribution potentielle de Gaston et de Peterson *et al.* (équation 1.1).

Nous notons alors que selon cette définition la niche potentielle comprend une partie dans \bar{M} qui possède les conditions abiotiques et biotiques requises par l'espèce. Cet espace écologique constitue une zone viable et donc sujet à être envahie si l'accès devient possible.

$$NP = A \cap B \quad (1.1)$$

1.3 Dynamiques des distributions d'espèces.

Nous avons vu précédemment qu'une niche peut exister ou être réalisée selon des conditions environnementales particulières. Or l'environnement n'est pas figé dans le temps, une modification soudaine de ce dernier contraint donc les espèces qui y vivent à s'y adapter. Cette modification peut être en faveur de l'espèce et lui permettre d'accéder à une plus grande part des ressources, ou au contraire la desservir; dans ce cas l'espèce devra se mouvoir dans des zones refuges, s'adapter (modifier ses préférences écologiques) ou disparaître. Les dynamiques des niches sont donc principalement dues aux interactions abiotiques et biotiques que les organismes ont avec leur environnement au cours du temps.

1.3.1 Interactions abiotiques

Naturellement, et si rien ne les contraint, les organismes d'une espèce vont occuper en priorité les lieux de leur niche fondamentale qui leur sont les plus propices (figure 1.4, la densité sera la plus haute pour les valeurs de *fitness* les plus élevées). Nous pouvons alors émettre l'hypothèse que des changements d'environnements abiotiques (changements climatiques, modifications géologiques du paysage, etc.) peuvent induire des migrations de populations vers les nouvelles conditions les plus favorables. Il se peut alors, que certains phénotypes soient privilégiés par rapport à d'autres. On assistera donc peu à peu à une adaptation, un décalage de la niche, qui peut mener à la création de nouvelles espèces. Les changements environnementaux peuvent également ériger des barrières, scindant les espèces en populations isolées qui, soumises à la dérive, divergent et forment de nouvelles espèces par spéciation allopatrique. Un changement de l'environnement physique peut provoquer la disparition d'une espèce, soit directement par diminution de sa niche fondamentale, soit indirectement parce qu'il affecte d'autres membres de l'écosystème dont l'espèce dépend : proies, prédateurs, commensaux ou compétiteurs. Ces interactions sont dites biotiques et sont détaillées dans la section suivante.

1.3.2 Interactions biotiques

La biogéographie d'une espèce (ou répartition géographique d'une espèce en fonction de son écologie) est dépendante du cortège d'espèces présent dans son écosystème. Ce phénomène est le résultat de modifications de la *fitness* de l'espèce par la présence d'autres organismes. L'exclusion compétitive [Gause, 1934; Grinnell, 1904] et le modèle proie-prédateur de Lotka-Volterra [Volterra, 1928] ont été parmi les premières interactions entre espèces modélisées par les écologues évolutionnistes.

L'étude de MacArthur [1958] sur les parulines (figure 1.7) illustre bien le principe d'exclusion compétitive. En effet, le scientifique a noté que la présence de certaines espèces contraint en lieu et en durée l'occupation d'un habitat ; en d'autres termes la niche réalisée s'exprime en fonction des relations inter-espèces et cela différemment, selon la cohorte d'espèces.

Par cet exemple MacArthur [1958] souligne que la coexistence entre espèces n'est possible que si ces dernières occupent des niches différentes, c'est à dire qu'elles se nourrissent ou qu'elles nichent dans le conifère en différents temps et lieux. Hutchinson [1957] signalait ce phénomène comme une réduction de similarité entre les niches d'espèces en compétition. Dans la figure 1.8 on peut voir que la *fitness* d'une espèce varie le long du gradient environnemental du fait qu'elle soit seule ou en présence d'autres espèces. La partie gauche de la figure représente la niche théorique de l'espèce isolée, soit sa niche fondamentale. La partie droite de la figure montre la niche en interaction avec les autres espèces de son

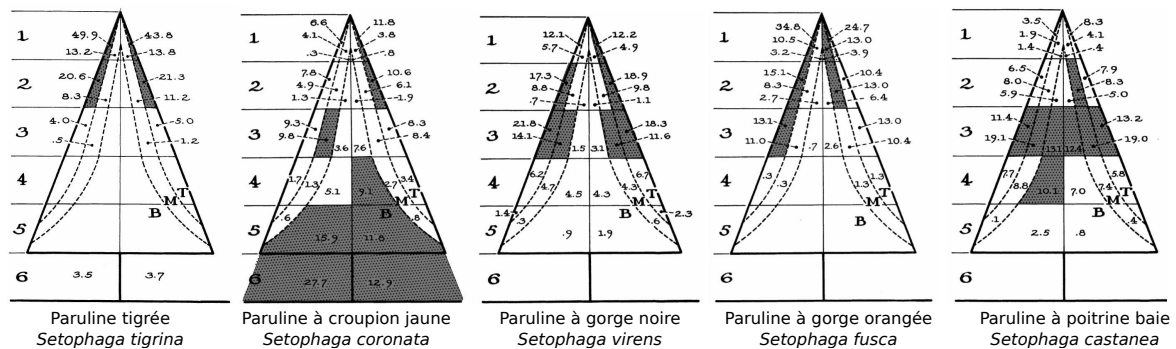


FIGURE 1.7 : Répartitions des niches écologiques de cinq parulines vivant dans les conifères. D'après [MacArthur \[1958\]](#). L'habitat : le conifère est schématiquement représenté par un triangle divisé par un axe suivant le tronc. Dans la partie de gauche de l'arbre sont indiquées des durées d'occupation de zone et dans la partie de droite des comptages de visites. La partie grisée correspond à la niche estimée de chaque espèce par rapport aux observations de l'auteur.

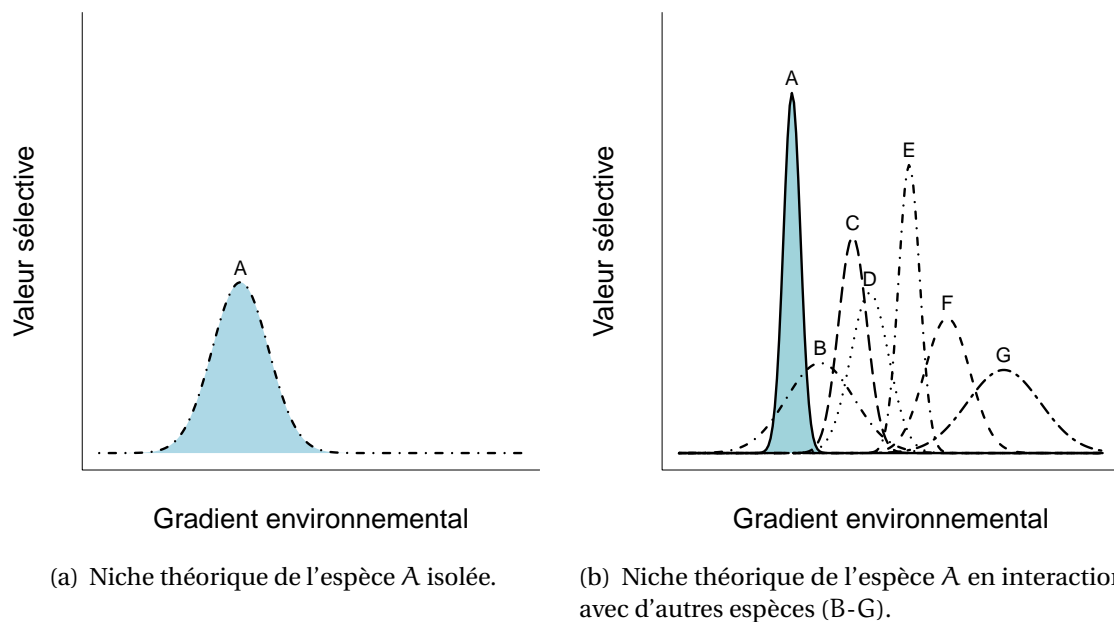


FIGURE 1.8 : Représentations théoriques de niches écologiques d'une espèce seule ou en communauté. On note que les interactions (de compétitions, de facilitations ou autres) entre les espèces ont provoqué un décalage de l'optimum de la niche ainsi qu'une baisse de son expression dans l'environnement (diminution de largeur). Ce schéma théorique a été construit sous l'hypothèse que toutes les espèces interagissent entre elles, soit directement comme A sur B (et réciproque), soit indirectement comme G sur A (et réciproque) par l'intermédiaire de toutes les autres.

environnement et correspond à la niche réalisée. On observe dans cette représentation théorique d'un écosystème⁶ des variations importantes de la niche réalisée :

- La distribution sur le gradient environnemental est comprise dans une gamme plus étroite lorsque l'espèce n'est plus isolée, signifiant que la largeur de la niche⁷ est affectée par les relations biotiques.
- L'optimum de la niche (situé au niveau du mode de la gaussienne) est la valeur du gradient environnemental pour laquelle la *fitness* de l'espèce est maximale. Sa valeur change en fonction de la présence d'autres espèces.
- L'aire sous la gaussienne a été fortement affectée par les interactions biotiques, suggérant que le succès global de la NR se trouve modifié par les interactions entre espèces.

La mise en évidence d'une répartition moins large le long du gradient environnemental, ainsi que les variations de *fitness*, illustrent un phénomène de spécialisation de niche pouvant conduire à une spéciation sur le long terme qui serait induite par des interactions biotiques.

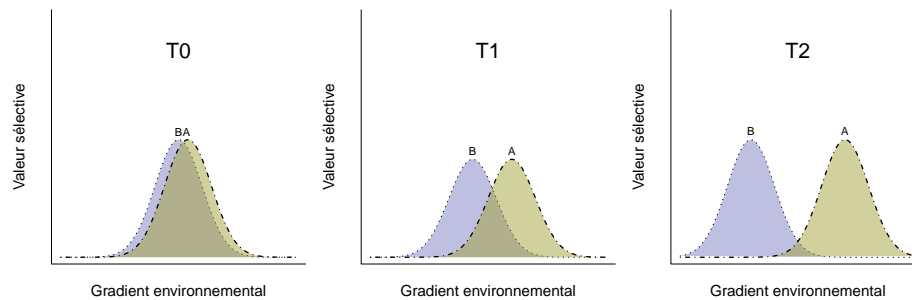
Le principe de Voltera-Gause⁸ [Gause, 1934], qui intègre des interactions biotiques de type proies-prédateurs et des notions de démographie, vient complexifier le concept hutchinsonnien et justifie la distinction entre niche fondamentale et niche réalisée. Il suggère que la distribution d'une espèce évolue dans le temps et dans l'espace du fait de ses interactions avec les environnements biotiques et abiotiques. Ces interactions conduisent les espèces à évoluer dans des milieux qui ne leur sont pas originellement optimaux, les poussant à s'adapter voir à modifier leur niche. On assiste alors dans les cas extrêmes à une spéciation sympatrique. Il existe des interactions biologiques négatives (compétition, prédation, parasitisme, amensalisme) comme des interactions positives (altruisme, commensalisme, mutualisme, symbiose). La figure 1.9 montre trois exemples de relations biotiques possibles conduisant à des décalages de niches menant à la spéciation (figure 1.9(a)), à l'extinction (figure 1.9(b)) ou à la symbiose (figure 1.9(c)). Dans un écosystème donné ces interactions se combinent pour finalement atteindre un équilibre. Ainsi chaque écosystème sera finalement constitué d'un assemblage d'espèces particulier, chacune réalisant une niche en accord avec celui-ci.

Dans la figure 1.10 nous avons schématisé les interactions biotiques et abiotiques d'un écosystème simple à deux facteurs environnementaux et trois espèces. Les niches fondamentales de chaque espèce peuvent être représentées par des ellipses dont les axes sont construits selon leur *fitness*. Chaque axe correspond à un facteur environnemental du plan.

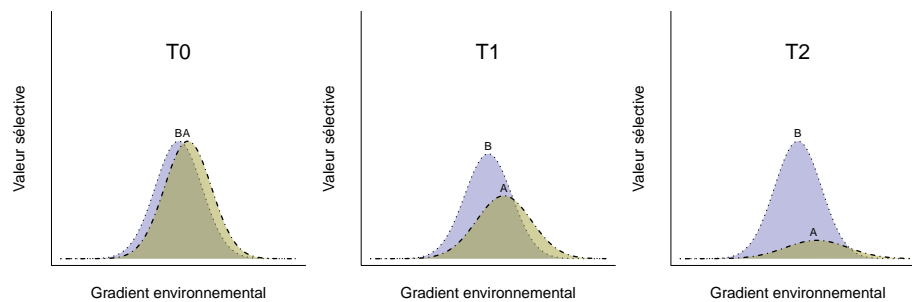
6. On obtiendrait de tels résultats en laboratoire si on connaissait toutes les composantes environnementales et toutes les espèces qui opèrent dans l'écosystème.

7. La largeur environnementale d'une niche correspond à la gamme du gradient environnemental que la niche peut occuper.

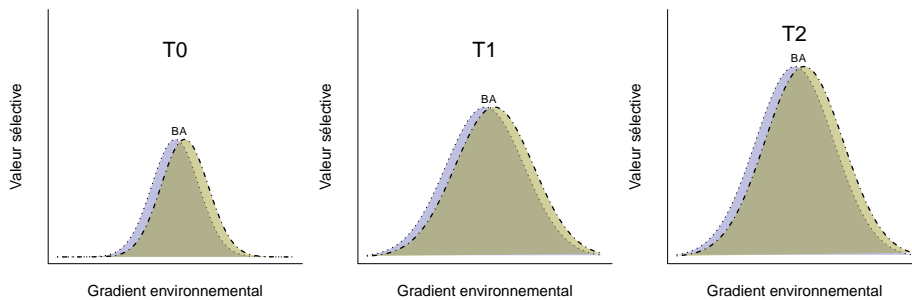
8. Aussi nommé principe de l'exclusion compétitive : deux espèces qui partagent et sont dépendantes de la même ressource ne peuvent pas coexister dans un même lieu pour une longue durée.



(a) Compétition/spéciation. Les espèces en compétition modifient leurs exigences environnementales, pour une coexistence à l'équilibre. Peu à peu les individus les plus adaptés sont sélectionnés. Dans un temps long ce phénomène peut aboutir à la création de nouvelles espèces.



(b) Compétition/extinction. Les espèces sont en compétition. L'une prend le pas sur l'autre (exclusion compétitive) jusqu'à extinction de cette dernière.



(c) Facilitation/extension. L'espèce A modifie l'environnement et le rend plus propice à B qui à son tour le rend plus propice à A. Dans un temps infini et en théorie se crée une symbiose entre A et B qui occuperont une partie de l'environnement plus grande que leurs niches fondamentales d'origine ne le permettaient.

FIGURE 1.9 : Évolution de deux niches relativement similaires au cours du temps. De la gauche vers la droite sont représentés différents états relatifs à la durée de l'interaction entre les deux espèces (A et B), T0 : pas d'interaction, T1 : situation intermédiaire, T2 : équilibre. Cette figure illustre deux exemples d'exclusion compétitive [Gause, 1934] et un exemple d'interaction positive [Callaway *et al.*, 1991].

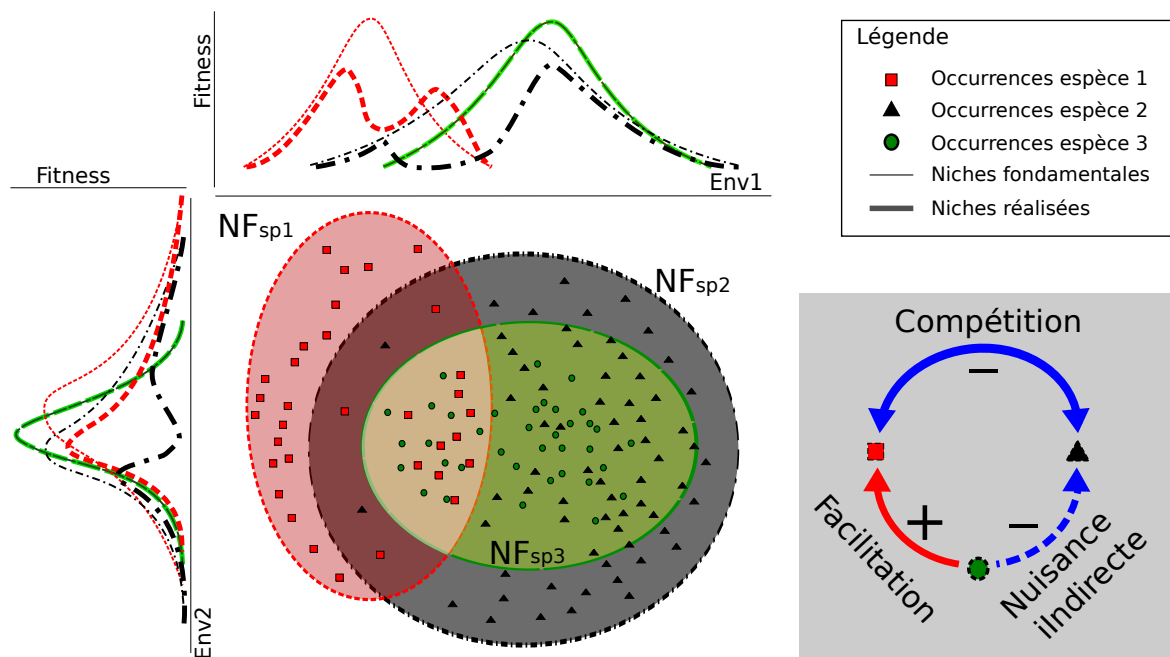


FIGURE 1.10 : Interactions biotiques entre trois espèces, représentativité des paramètres abiotiques (de la niche fondamentale [NF]) par les occurrences d'espèces en interactions. Les ellipses représentent les niches fondamentales des espèces et les différents symboles les occurrences. Dans le cadre grisé sont schématisées les interactions inter-espèces : l'espèce 1 (sp1) et l'espèce 2 (sp2) sont en compétition et l'espèce 3 (sp3) facilite sp1. Remarquons qu'il existe aussi des interactions indirectes, par exemple : la facilitation causée par sp3 confère un avantage à sp1 sur sp2. Les courbes en traits fins représentent les *fitness* théoriques (selon les NF) des espèces en fonction des gradients environnementaux (Env1 et Env2), alors que les courbes en gras représentent les niches réalisées basées sur les occurrences le long des mêmes gradients.

Les trois espèces sont liées par les interactions suivantes : sp1 et sp2 sont en compétition et sp1 est facilitée par la présence de sp3. Cette figure met en évidence, par l'observation des courbes des niches réalisées et la distribution des occurrences dans l'espace écologique, les interactions indirectes des espèces entre elles. En effet, isolément, sp2 et sp3 n'interagissent pas. En revanche, lorsque les trois espèces sont présentes on note un effet négatif indirect de sp3 sur sp2 du fait de l'avantage qu'il a été conféré à sp1 par la facilitation.

Ces interactions biotiques directes et indirectes nous conduisent à une vision différente de l'écologie et à reconsidérer la position de la niche écologique dans l'environnement en fonction des communautés y demeurant.

1.3.3 Vision moderne de la niche écologique

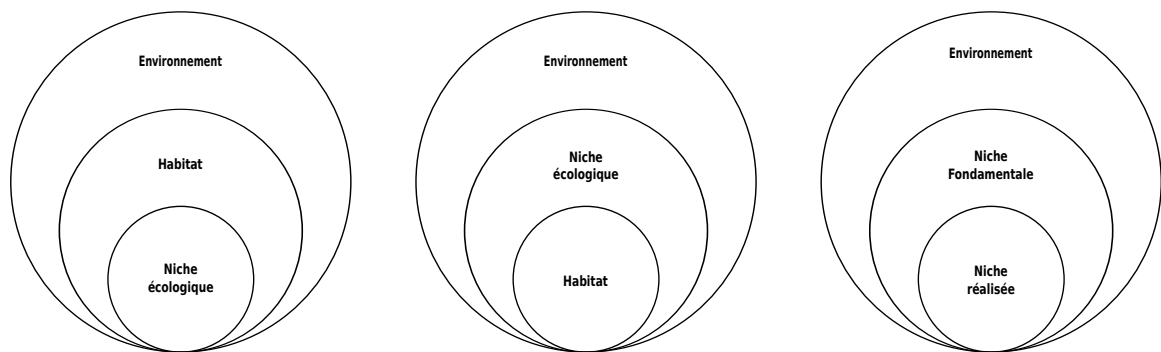
Du fait de la découverte antérieure des interactions biotiques négatives de type compétition et prédation, la théorie des niches (voir section 1.1) est principalement axée sur ces dernières. Elle vise à mesurer en appliquant les principes de Lotka-Volterra le degré de compétition entre espèces pour connaître leur abondance et estimer la densité des individus. Pulliam [2000] aborde le concept de niche de façon moins dynamique, donnant moins de poids à la compétition et intégrant les processus stochastiques dus aux interactions avec toutes les espèces de la communauté. Ce concept accorde également de l'importance aux déplacements de l'espèce [Araújo *et al.*, 2005; Guisan et Thuiller, 2005; Peterson, 2006; Pulliam, 2000; Soberon et Peterson, 2005; Svenning et Skov, 2004], à sa démographie, aux aléas de son histoire [Brown, 1995; Soberon et Peterson, 2005], aux espèces constructrices de leur environnement⁹, ainsi qu'à la dynamique sources-puits. Plus récemment, Chase et Leibold [2003] ont mis à jour la définition de la niche écologique comme étant un espace regroupant les conditions nécessaires permettant à l'espèce d'avoir un taux de croissance positif, c'est-à-dire un taux de natalité supérieur au taux de mortalité. La conception moderne de la niche grinellienne de ces auteurs est celle qui convient le mieux aux ENM.

1.3.4 Niche ou habitat deux ensembles non comparables.

Comme nous l'avons vu dans les sections précédentes, l'habitat et la niche écologique sont deux divisions de l'environnement ne relevant pas des mêmes critères et ne fournissant pas la même information. Il est donc délicat de les comparer. Les deux premiers cas de la figure 1.11 montrent l'agencement entre niche écologique et habitat selon deux définitions de l'habitat : 1.11(a) à l'échelle du paysage (Niche écologique \subset Habitat) et 1.11(b) où la description du milieu est bien plus spécifique (Habitat \subset Niche écologique).

Souvent évaluée uniquement à partir de la répartition géographique des occurrences, la distribution géographique peut délimiter par conséquent une aire incluant des régions dont les conditions environnementales vont à l'encontre de l'écologie de l'espèce. Les deux autres caractéristiques présentées ici, l'habitat et la niche écologique, tiennent compte de l'environnement dans lequel l'espèce évolue. Les positions relatives entre celles-ci dans l'espace environnemental présenté figure 1.11(a) et figure 1.11(b), sont une vision simpliste et quelque peu erronée des emboîtements de ces ensembles écologiques (encore nommés espaces ou compartiments écologiques). En effet l'habitat et la niche sont décrits selon des critères bien distincts, par analogie l'habitat équivaldrait à l'« adresse » de l'espèce alors que sa niche serait sa « profession » [Odum, 1959]. Les notions d'habitat et de niche écologique, respectivement définies dans la section 1.2.2 et la section 1.2.3, apportent des informations complémentaires quant à l'écologie de l'espèce, son rôle et sa place dans l'environnement.

9. L'auteur pointe les interactions réciproques entre l'espèce et son environnement : l'espèce s'adapte à son environnement mais aussi le modifie, se créant ainsi sa propre dynamique et influençant également celle des autres espèces.



(a) Cas où l'habitat de l'espèce est décrit par un élément commun du paysage.

(b) Cas où l'habitat décrit très spécifiquement le milieu de vie de l'espèce.

(c) Cas où la niche écologique est compartimentée en deux : niche fondamentale et niche réalisée.

FIGURE 1.11 : Différentes conceptions d'enchevêtrement des ensembles écologiques.

Les définitions de ces ensembles restent encore quelque peu ambiguës et bien souvent en ENM il est difficile de les différencier l'un l'autre car des données relevant de chacun de ces ensembles ont pu être utilisées pour établir les modèles.

1.4 Modélisation de niche ou d'habitat, que modélise-t-on ?

Il est difficile de comparer l'habitat et la niche écologique d'une espèce d'autant plus que nous avons vu que cette niche pouvait être compartimentée en niche fondamentale et niche réalisée (section 1.2.3). Bien que les occurrences d'espèces soient le reflet de l'habitat nous allons voir dans les parties suivantes que les modélisations de niches (ENM) ne parviennent pas pour autant à reconstituer l'habitat strict ou la niche réalisée avec un fort degré de fiabilité.

Outre les définitions ambiguës au niveau des compartiments écologiques (niche écologique ou habitat), existe également un débat au sein des modélisateurs de niches écologiques [Araújo et Guisan, 2006; Kearney, 2006]. S'ils s'appuient tous sur la définition de la niche écologique d'Hutchinson [1957] les avis divergent concernant la partie de la niche réellement modélisée [Araújo et Guisan, 2006]. Ces auteurs se questionnent notamment sur l'influence des espèces environnantes dans les modélisations de niches.

1.4.1 Impact des interactions biotiques sous-jacentes dans l'élaboration des modèles

Après l'obtention d'un modèle la question est de savoir si la niche modélisée est le reflet d'interactions biotiques, c'est-à-dire si elle décrit mieux la niche réalisée que la niche fondamentale.

La présence d'une espèce est dépendante non seulement de son environnement abiotique et du milieu qui lui est accessible mais elle est aussi influencée par d'autres espèces de son écosystème. De ce fait, nous pourrions penser que toutes modélisations utilisant des occurrences basées sur des observations en milieu naturel intègrent donc indirectement une dimension biotique (Hypothèse du bruit Eltonien, ENH¹⁰) [Soberón et Nakamura, 2009]). Les avis des modélisateurs sont partagés quant à la nature exacte des niches modélisées en ENM. Pour construire un modèle de niche écologique, deux types de données sont nécessaires : des couches de données environnementales d'une part et des données d'occurrences d'autre part (voir section 2.2 pour plus de détails). Dès lors, deux questions concernant la part des interactions biotiques dans la modélisation peuvent être soulevées :

- Quelle est l'intensité du reflet des interactions biotiques parmi les occurrences et quel en est l'impact sur la prédiction de niche ?
- Les variables environnementales, si elles sont suffisamment exhaustives, peuvent-elles refléter indirectement les interactions biotiques ?

Des auteurs estiment que la reconstruction de niches à partir de données abiotiques ne peut que conduire à l'obtention de modèles approximant la niche fondamentale [Araújo et Guisan, 2006; Peterson et Cohoon, 1999; Soberon et Peterson, 2005]. D'autres soutiennent que les modèles reflètent l'expression géographique de la niche réalisée [Austin *et al.*, 1990; Guisan et Zimmermann, 2000; Pearson et Dawson, 2003 cités par Araújo et Guisan, 2006]. Les défenseurs de la première hypothèse soulignent l'influence de la résolution des études souvent menées à échelles régionales ou continentales. Celle-ci masquerait l'effet des interactions biotiques locales selon le même principe que nous évoquons section 1.2.1 [Araújo *et al.*, 2005; Guisan et Thuiller, 2005; Peterson, 2006; Pulliam, 2000; Soberón, 2007; Soberon et Peterson, 2005; Svenning et Skov, 2004]. De cette manière, ces auteurs invoquent principalement deux arguments pour montrer que les interactions biotiques ne peuvent pas être perçues lors d'une modélisation à grande échelle. Ils énoncent premièrement, que les limites de dispersions peuvent être un facteur plus important que la compétition pour expliquer l'absence d'une espèce, ou peuvent au contraire expliquer sa cooccurrence du fait que les espèces peuvent se déplacer et éviter la concurrence à l'accès aux ressources ; deuxièmement, que le maillage des modèles à grande échelle, pouvant atteindre plusieurs dizaines de kilomètres, biaise la perception des phénomènes d'exclusions compétitives ou de cooccurrences liées aux équilibres locaux. On peut ajouter à cela le fait que des espèces dont les niches se chevauchent peuvent coexister dans une même zone géographique pour

10. sigle anglais pour *Eltonian Noise Hypothesis*

peu qu'elles soient faiblement compétitives entre elles ou bien qu'elles n'occupent pas la niche ou la portion de niche concernée au même instant ([Amarasekare, 2003] dans [Araùjo et Guisan, 2006]). Dans ces situations les interactions biotiques (principalement la compétition) ne seraient pas prises en compte dans la modélisation et la niche reconstruite équivaldrait à une approximation de la niche fondamentale.

Araùjo *et al.* [2014]; Soberón et Nakamura [2009] suggèrent que la résolution des couches environnementales (voir section 2.3.1.3) peut influencer le résultat dans le sens contraire. Bien que les facteurs eltoniens étroitement corrélés aux variables abiotiques permettent de capturer une part importante de la signature biotique [Brewer et Gaston, 2003; Elith et Leathwick, 2009], la taille du grain ¹¹ peut amplifier certaines interactions biotiques locales. En effet, en un point précis l'espèce peut être présente du fait des interactions biotiques ce qui assigne une valeur de présence pour le carré entier où se trouve cette occurrence, quelles que soient les conditions environnementales moyennes de ce dernier [Prinzing *et al.*, 2002]. Le modèle considère alors ces conditions comme étant favorables à l'espèce et les incorpore à la niche.

Nous voyons ainsi que, en dépit de la sensibilité des algorithmes au nombre d'occurrences (voir section 2.3.1.1), l'effort d'échantillonnage ¹² doit être dans certains cas modéré en rapport avec la résolution du maillage car un échantillonnage faible avec une haute résolution spatiale ne décrit pas correctement la niche réalisée et à l'inverse un échantillonnage trop intense avec une résolution grossière produit une modélisation en désaccord avec la niche fondamentale.

Araùjo *et al.* [2014] testent l'hypothèse du bruit Eltonien [Soberón et Nakamura, 2009].

D'après ces auteurs l'ENH implique que :

- les interactions biotiques jouent un rôle plus important à petite échelle qu'à large échelle ;
- une modélisation intégrant des facteurs biotiques apportent plus d'information qu'une modélisation n'intégrant que des facteurs abiotiques ;
- l'on observe des différences plus marquées pour les espèces ayant des interactions biotiques fortes lorsque l'on compare les niches prédites avec ou sans facteurs biotiques.

Ces auteurs ont mené une étude sur des perroquets qui leur a permis de rejeter l'ENH car les résultats étaient en désaccord avec les implications citées précédemment. Cependant on peut expliquer ce résultat par une forte corrélation entre les variables biotiques et abiotiques que ces derniers utilisent. En effet, si dans un premier temps les auteurs prennent le soin de réduire la redondance d'information en éliminant les variables abiotiques corrélées entre elles (comme préconisé par [Johnson *et al.*, 2002]), ils la favorisent quand ils incluent les variables biotiques issues de modélisations à partir des variables abiotiques d'origine. Malgré

11. On parle de grain grossier pour une faible résolution et de grain fin pour une résolution élevée.

12. Nombre minimum d'occurrences représentatif.

ce biais, cette étude met en avant les apports de variables biotiques, qui se trouvent être en réalité des combinaisons de variables abiotiques, à la modélisation. On en déduit que les résultats d'une modélisation sont plus pertinents lorsque l'information environnementale est plus détaillée, c'est-à-dire plus les variables environnementales sont diverses et variées, car elles permettent indirectement de déchiffrer l'environnement biotique.

En théorie une méthode de modélisation parfaite pourrait se contenter seulement de quelques variables abiotiques fondamentales dont dépendraient non seulement toutes les autres variables abiotiques mais aussi toutes les interactions biotiques. Or cela impliquerait des combinaisons très complexes. C'est pourquoi le fait d'ajouter aux modélisations des variables supplémentaires ne peut qu'affiner les résultats car les algorithmes prédictifs peuvent ainsi faire émerger des relations insoupçonnées. D'autres exemples prouvent que l'inclusion de variables biotiques modifie significativement et positivement les résultats de modélisation de niche écologique [Araújo et Luoto, 2007; Heikkinen *et al.*, 2007; Prinzing *et al.*, 2002].

La représentativité des interactions biotiques dans les occurrences et leur impact sur la modélisation nous conduit alors à nous interroger sur la niche réellement modélisée.

1.4.2 L'espace écologique modélisé

Du fait des ambiguïtés énumérées précédemment, il est difficile de différencier avec certitude, si les compartiments écologiques modélisés en SDM et ENM relèvent plus de l'habitat ou de la niche écologique. Cependant nous avons vu que le concept d'habitat n'était pas toujours clair et que sa pertinence était relative au degré de précision de sa description. C'est pourquoi, en ENM, il vaut mieux privilégier la compartimentation de l'environnement pouvant être occupé par l'espèce en différentes niches (figure 1.11(c)), qui plus est peuvent être évaluées qualitativement et quantitativement. Dans la suite du texte, et dans les figures, l'abréviation NM fera référence au résultat de la modélisation, c'est-à-dire à la niche écologique prédite.

Afin de savoir quelle partie de la niche est susceptible d'être modélisée, observons de plus près les deux méthodes généralement utilisées [Kearney, 2006] afin d'estimer la distribution potentielle d'une espèce. La première est une approche mécaniste qui consiste à évaluer au préalable la niche écologique de l'espèce, en essayant d'être le plus précis possible. Cette opération peut s'effectuer en laboratoire ou être fondée sur des observations *in situ* de l'organisme. Une fois la niche établie la distribution est estimée en recherchant spatialement tous les lieux qui remplissent les exigences environnementales de celle-ci. La deuxième technique est dite corrélative (figure 1.12) et consiste à déterminer l'écologie de l'organisme par des analyses de similitudes entre des données d'occurrences et environnementales. La distribution est ensuite estimée par projection spatiale, comme pour l'approche mécaniste.

Les deux types de modélisation conduisent à des résultats différents car il n'intègrent pas le même type de données. L'approche mécaniste est une application directe de la formule

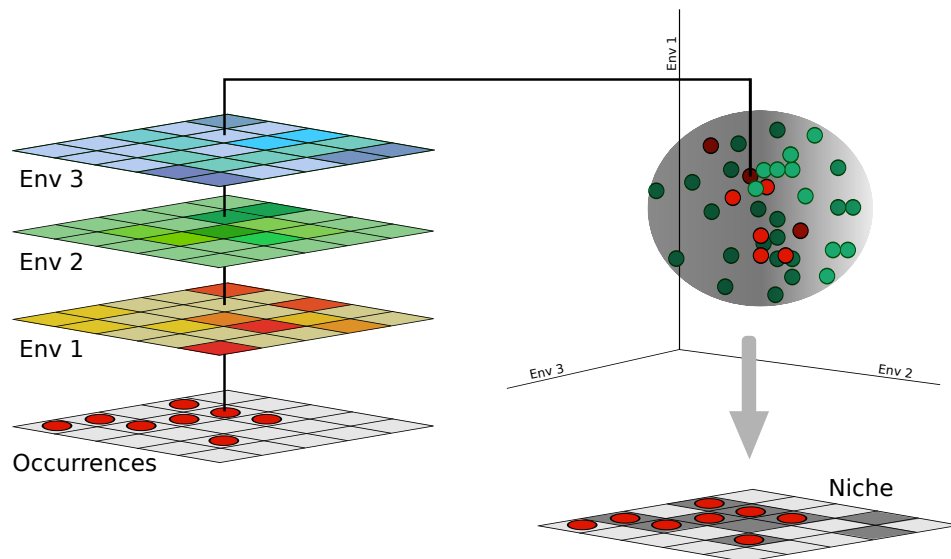
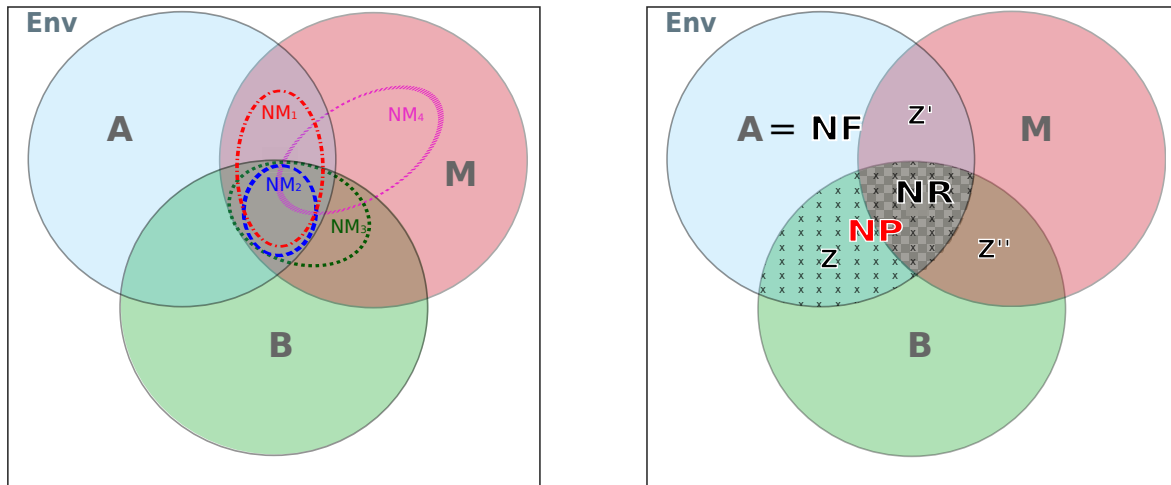


FIGURE 1.12 : Modélisation de niche par approche corrélative. Exemple de la reconstruction d'une niche à trois dimensions environnementales et sa projection géographique. Les points rouges représentent les occurrences. Chacune des couches de données environnementales (Env1, Env2 et Env3) possède une valeur numérique de la variable qu'elle représente en tout point de grille. Les valeurs qui correspondent aux occurrences sont reportées sur les axes du graphique de droite. Les corrélations entre les variables environnementales et les occurrences sont calculées puis des occurrences sont simulées (points verts). Le volume en gris symbolise l'ensemble des occurrences possibles, soit la niche modélisée. Enfin la niche est projetée sur une carte où les prédictions de présence sont représentées par des cases grisées.

de l'hypervolume d'[Hutchinson](#). Dans ce cas la niche modélisée est la niche fondamentale $NM = NF$. Toutefois la prédiction peut être affinée en retirant *a posteriori* la zone inaccessible à l'espèce (correspondant au \bar{M} du diagramme BAM, figure 1.6). La niche modélisée obtenue correspond alors à l'intersection de A et de M soit $NM = A \cap M$ (exemple : NM1 de la figure 1.13(a)). De la même manière les interactions biotiques, négatives comme positives, peuvent être intégrées à la modélisation en superposant les distributions connues des espèces concernées et en les appliquant comme des filtres sur le modèle prédit (par exemple, en diminuant la probabilité de présence dans les lieux où un compétiteur est recensé). Cette fois la niche modélisée, toujours incluse dans $A \cap M$, sera réduite et tendra à égaler l'intersection des ensembles B , A et M ($NM \rightarrow B \cap A \cap M$, exemple : NM2 de la figure 1.13(a)). Il faut préciser toutefois que cette approche reste difficile à mettre en place pour certaines espèces, notamment les disparues, que les relations biotiques sont parfois difficiles à identifier, ainsi que la difficulté de reproduire certains facteurs environnementaux en laboratoire.

L'approche corrélative consiste à prélever des occurrences (x) dans le milieu naturel



(a) Trois exemples de niches modélisées à partir de divers lots d'occurrences dont les représentations des interactions biotiques divergent. NM₁ est une niche modélisée à partir d'occurrences minimisant les interactions négatives. NM₂ est une niche modélisée à partir de données biotiques bien représentées. NM₃ est un modèle créé à partir de données où la facilitation est sur-représentée. NM₄ est un modèle où les occurrences sélectionnées ne sont pas représentatives des interactions avec l'environnement.

(b) Trois zones, recoupant chacune deux des trois ensembles seulement, se distinguent : $Z = B \cap A \cap \overline{NR}$, zone qui peut être potentiellement envahie [Peterson *et al.*, 2011]; $Z' = A \cap M \cap \overline{NR}$, où l'environnement biotique n'est pas favorable; et $Z'' = B \cap M \cap \overline{NR}$, hors de la niche fondamentale.

FIGURE 1.13 : Positions des niches dans le diagramme BAM. La niche fondamentale : $NF = A$; la niche potentielle : $NP (A \cap B)$, zone représentée par des croix; l'aire de distribution géographique de l'espèce, que l'on peut assimiler à la niche réalisée : NR (zone en damier, $B \cap A \cap M$).

de l'espèce. Celles-ci sont recueillies à la fin du processus évolutif de mise en place de l'espèce dans son environnement, et sont donc en théorie comprises dans la niche réalisée restreinte à l'espace accessible à l'espèce, soit incluse dans NR . On peut donc considérer les occurrences comme le reflet d'interactions biotiques et abiotiques ($B \cap A$) [Kearney, 2006] exprimées dans le milieu M qui leur est accessible (soit $x \in B \cap A \cap M$).

Les données abiotiques, correspondant aux occurrences utilisées pour établir le modèle, ne sont donc pas représentatives de la niche fondamentale mais sont fortement influencées par les niches fondamentales des espèces qui interagissent avec l'espèce étudiée. Ainsi à partir d'un jeu d'occurrences représentatif de l'écologie de l'espèce on obtient une niche modélisée correspondant à l'intersection des environnements biotiques et abiotiques accessibles $NM = B \cap A \cap M$, soit $NM = NR$. Or suite aux aléas de l'histoire de l'espèce, au biais dus à la résolution de l'analyse et à la dynamique « source-puits », les occurrences ne se situent pas exclusivement dans NR . Gaston [2003]; Pulliam [2000] suggèrent que les lieux

géographiquement et environnementalement proches de NR (zone des sources), peuvent accueillir l'espèce malgré des conditions qui ne lui sont pas tout à fait favorables. Cette zone d'accueil opportuniste est nommée zone des « puits ». Elle correspond à $\overline{NR} \cap M$ de la figure 1.6 où les absences, représentées par les cercles vides, pourraient aussi bien indiquer de telles occurrences. Néanmoins, la probabilité d'existence de « puits » augmente en fonction de la proximité environnementale et spatiale des sources. Ceci suggère que la probabilité de telles occurrences est plus grande dans les zones Z' et Z'' de la figure 1.13(b) car les « puits » se situent dans ces cas à l'intersection de deux ensembles définissant l'écologie de l'espèce. Selon la représentativité des interactions biotiques et abiotiques des occurrences échantillonnées pour construire le modèle les niches obtenues seront comprises dans M avec une probabilité forte d'émarger dans NR, un peu plus faible dans Z' et Z'' , et assez faible mais non nulle d'être dans $M \cap \overline{A} \cap \overline{B}$ (équation 1.2). Trois exemples de niches modélisées possibles sont représentées figure 1.13(a).

Pour reconstruire des niches écoculturelles nous nous servons des techniques utilisées pour la modélisation de niches écologiques. L'ENM est une approche corrélative. Il est donc important de connaître les compartiments écologiques pouvant être modélisés par cette approche et de les discerner afin d'interpréter correctement les résultats obtenus. La figure 1.13(b) montre trois ensembles qui sont en dehors de la niche réalisée (NR) de l'espèce mais pouvant cependant être prédits quand les paramètres de la modélisation ne sont pas totalement contrôlés, ou que les occurrences échantillonnées ne sont pas représentatives de la niche réalisée. Ces ensembles sont Z lorsque le milieu accessible (M) n'est pas estimé correctement, Z' lorsque les interactions biotiques (B) sont mal connues et Z'' quand certains facteurs abiotiques (A) ne sont pas considérés. Ces zones écologiques ont chacune leurs propres caractéristiques :

- Z est une zone qui peut être potentiellement envahie [Peterson *et al.*, 2011] car les conditions abiotiques et biotiques requises au développement de l'espèce sont réunies. Cependant elle ne peut pas l'atteindre. On peut imaginer divers *scenarii* pour la rendre potentiellement accessible comme le développement d'une espèce facilitatrice, un événement géologique permettant l'accès ou l'importation fortuite de l'espèce dans cette zone jusque là inatteignable.
- Z' est une zone de la niche fondamentale qui se trouve dans le milieu accessible. Elle présente donc des conditions abiotiques viables, mais les conditions biotiques y régnant sont défavorables. Cette zone pourrait être colonisée par l'espèce si les interactions avec les autres organismes venaient à changer, par exemple avec l'extinction d'une espèce prédatrice ou l'introduction d'une espèce facilitatrice.
- Z'' est une zone hors de la niche fondamentale. Théoriquement l'espèce ne peut pas y vivre, mais l'interaction avec d'autres espèces, notamment les facilitatrices, peut recréer localement les conditions de la NF en jouant le rôle de microhabitats ou en modifiant la nature du milieu.

En ENM l'espace accessible à l'espèce est déterminé en avance et est appliqué comme un masque sur les modèles de niches prédits. Cela a pour effet de contraindre les prédictions à ne pas déborder hors de la zone qui correspond à l'expression géographique M . Ceci signifie que la zone Z se trouve toujours exclue des prédictions car elle se situe dans \overline{M} . En revanche les prédictions peuvent se trouver dans les deux autres zones Z' et Z'' , car la position de la niche modélisée de l'espèce dans l'espace écologique est fonction de la représentativité des interactions biotiques parmi les occurrences échantillonnées. Ainsi selon cette représentativité, la position du centre d'une niche dans NR varie¹³ et la niche modélisée débordé plus ou moins dans les zones Z' et Z'' . On remarquera l'importance de bien définir¹⁴ le M de l'espèce car ce choix exclut une partie (Z) de la niche potentielle (NP , définie par Gaston [2003]; Peterson *et al.* [2011]; Soberón et Nakamura [2009]). Cela nous amène à définir un espace de modélisation potentiel que nous appellerons niche modélisée potentielle (NMP). Ce nouvel espace qui représente les prédictions de présence potentielle est inclus dans M . Or ces prédictions y sont inégalement réparties car les probabilités de prédictions sont corrélées aux probabilités des zones pouvant jouer le rôle de « puits ». Conséquemment, les prédictions ont une forte probabilité d'exister dans NR , moins forte dans Z' et Z'' et plus faible, mais non nulle, dans le reste de la zone accessible M (voir équation 1.2).

$$\forall x \in NMP, \quad p_{x \in NR} > p_{x \in Z' \mid x \in Z''} > p_{x \in (M \cap \overline{NR})} \quad (1.2)$$

On peut considérer non significative, toute NMP qui se trouve dans l'espace accessible mais hors des environnements biotiques et abiotiques qui conviennent à l'espèce. Cette probabilité $p_{x \in (M \cap \overline{A \cup B})}$ est théoriquement presque nulle car il s'agirait de « puits » très éloignés des conditions optimales de la réalisation de la niche de l'espèce. Ainsi NMP devient égale à l'union de la niche réalisée (NR) et des zones Z' et Z'' . (équation 1.3).

$$NMP = NR \cap Z' \cap Z'' = (A \cap M) \cup (M \cap B) \quad (1.3)$$

Le tableau 1.1 montre les espaces écologiques modélisés de NMP en fonction de la représentativité de l'écologie de l'espèce dans l'échantillonnage (interactions biotiques et abiotiques sur-représentées ou sous-représentées) ainsi que la proportion des occurrences parmi les « puits » ou les « sources » utilisées. On voit dans ce tableau qu'une niche modélisée égale à NR serait issue d'occurrences représentant parfaitement l'écologie de l'espèce ; qu'une niche modélisée débordant dans Z' équivaldrait à une sur-représentation des interactions biotiques négatives et des interactions abiotiques positives, ou un mélange de ces interactions ; et qu'un modèle qui émerge largement dans $\overline{NR} \cap M$ serait dû

13. Pour des occurrences bien représentatives de l'écologie de l'espèce le centre d'une niche modélisée se trouvent proches du centre de NR .

14. Barve *et al.* [2011] montre également les variations que confèrent différentes définitions du M pour la modélisation d'une même espèce sous les mêmes conditions environnementales

Tableau 1.1 : Niches modélisées en fonction de la représentativité de la niche réalisée par les occurrences échantillonnées. Les niches modélisées (NM1, NM2, NM3 et NM4) couvrent les régions NR, Z', Z'' et le reste de M. Les symboles ">0" et "<0", indiquent les natures positives ou négatives des interactions biotiques et abiotiques. Le signe "+", signifie que la niche modélisée représente exagérément l'interaction concernée, "-" une interaction sous-représentée par le modèle et "?" celles représentées de manière non déterminée. Le signe "*" signifie que les occurrences sont représentatives de l'écologie de l'espèce, lorsque l'ensemble des occurrences représente l'intégralité des interactions. Nous sommes alors dans le cas idéal où la niche modélisée est proche de la niche réalisée.

Niche modélisée	Compartiment dominant	Interactions biotiques		Interactions abiotiques		Nature des occurrences	
		>0	<0	>0	<0	Puits	Sources
NM2	NR	*	*	*	*	*	*
NM1	Z'	-	+	+	?	+	-
NM3	Z''	+	?	-	+	+	-
NM4	$M \cap \overline{A \cup B}$?	-	?	-	+	-

à une sur-représentation des occurrences issues des puits. Le résultat d'une modélisation faite à partir d'occurrences qui représentent une partie restreinte de l'écologie de l'espèce, par exemple prélevées dans l'optimum de la niche, peut aussi se traduire par une niche modélisée n'intégrant pas une partie de NR (exemple NM2 figure 1.13(a)). En effet, une sur-représentativité des interactions biotiques négatives ou une sous-représentation des interactions biotiques positives confinent les prédictions dans un espace écologique plus petit qu'il ne l'est en réalité. Cette erreur d'omission, peut également être due à un sur échantillonnage dans les conditions abiotiques les plus favorables ou à une sous-estimation de la mobilité.

La représentativité de l'écologie de l'espèce parmi les occurrences joue donc un rôle sur la proportion qu'occupe la niche modélisée potentielle (NMP) dans les ensembles concernés. Une bonne connaissance de l'écologie de l'espèce est nécessaire avant la modélisation de niche afin d'utiliser les paramètres environnementaux adéquats et de sélectionner les occurrences correctes. Dans certains cas ces choix n'étant pas intuitifs, on élimine *a posteriori* les prédictions issues des occurrences marginales en appliquant un seuil [Liu *et al.*, 2005; Peterson *et al.*, 2007] (voir section 7.1).

1.5 Évolution constante des niches écologiques

Les espèces sont en constante évolution (l'hypothèse de la reine rouge [Van Valen, 1977]). Or une reconstruction de niche écologique est une image à un temps précis de la répartition d'une espèce qui résulte de son histoire évolutive. Nous avons vu précédemment que la spéciation pouvait être induite par le biais d'interactions biotiques, de changement environnementaux ou de processus stochastiques. La spéciation peut intervenir de façon

soudaine par la voie de la sélection naturelle ou plus lentement suivant des ségrégations allopatriques ou sympatriques. Une ENM conduite pour différents climats (époques ou lieux différents) permet d'apercevoir, soit une réduction, une expansion ou un décalage de niches [Crisp *et al.*, 2009; Losos, 2008; Nogués-Bravo, 2009; Warren *et al.*, 2008; Wiens, 2004], soit une conservation de niche [Peterson, 2011; Peterson *et al.*, 1999]. L'hypothèse de la conservation de niche [Harvey et Pagel, 1991 cités par Wiens *et al.*, 2010], explique la géographie d'une espèce par l'expression d'un caractère hérité. La conservation de niche qui apparaît comme contradictoire avec la notion d'évolution est un processus d'adaptation qui offre l'opportunité à une espèce de puiser dans son patrimoine génétique des réponses à de nouveaux challenges environnementaux.

Nos travaux consistent à appliquer un raisonnement similaire pour étudier les relations cultures/environnements et apprécier l'évolution de celles-ci dans différents cadres environnementaux. Pour cela nous ne nous plaçons plus au niveau de la niche écologique mais au niveau de la niche écoculturelle.

1.6 Concept de niches écoculturelles

En 2006, Banks *et al.* introduisent le concept de niche écoculturelle (ECN¹⁵). Ces auteurs démontrent qu'il est possible de modéliser des relations cultures/environnements grâce à l'emploi d'outils que les écologues utilisent pour modéliser les niches écologiques d'espèces animales ou végétales. En rapport avec l'ENM les auteurs ont nommé cette technique *Eco-Cultural Niche Modeling* (ECNM).

1.6.1 Définition de la niche écoculturelle

Les groupes humains évoluent dans des territoires sociaux [Clark, 1975 cité par Banks *et al.*, 2011] au sein desquels ils mènent des activités caractéristiques. La cohésion de ces groupes peut être renforcée par des liens de parentés, des liens sociaux, une langue commune, une culture matérielle propre et des droits d'accès exclusifs au territoire [Verhart, 1990]. L'ensemble des systèmes technologiques, des modes d'occupation d'un territoire et des stratégies de subsistance, partagés et transmis par une population culturellement cohésive, dans un cadre paléoenvironnemental donné est appelé système adaptatif. L'ECNM permet de corréliser ces systèmes adaptatifs avec leur environnement offrant ainsi la possibilité d'identifier les processus culturels sous-jacents [Banks, 2013].

Une ECN peut donc être définie comme l'ensemble des lieux possédant des caractéristiques environnementales proches, où des traits culturels comme une langue commune, une même économie, un même art mobilier, une technologie similaire ou des modes de vie très proches ont la possibilité de s'exprimer et de se transmettre [Banks *et al.*, 2011, 2006].

15. Eco-Cultural Niche

1.6.2 Principe de l'ECNM

La modélisation de niches écoculturelles est une approche dérivée des sciences de la biodiversité. Elle consiste à établir les relations entre traits culturels et environnements ainsi qu'à déterminer les facteurs environnementaux les plus influents. Appliquée à des populations du passé, cette modélisation consiste, par exemple, à reconstituer des réponses culturelles, des variations de mode de subsistances ou de systèmes techniques qui ont eu lieu face à des changements environnementaux. Afin d'être validés, les résultats de la modélisation sont alors confrontés, dans la mesure du possible, à la réalité archéologique.

L'ECNM se réalise via les mêmes outils que l'ENM, c'est-à-dire à l'aide d'algorithmes prédictifs de niches écologiques (présentés dans le chapitre 2). Cette modélisation nécessite le même type de données que l'ENM et repose sur les mêmes principes. Elle possède donc les mêmes atouts et les mêmes biais. Ainsi la lecture et l'interprétation des niches écoculturelles se fait de la même manière que celle des niches écologiques et les mêmes mécanismes évolutifs peuvent être mis en évidence. La pertinence des ECN prédites est, comme en ENM, dépendante des points d'occurrences échantillonnés ainsi que de la résolution des données environnementales.

Notons qu'à la différence de l'ENM les interactions biotiques dans l'ECNM comprennent, en plus des interactions biotiques classiques, les échanges culturels, les barrières culturelles et les phénomènes d'acculturation qui peuvent s'exercer sur de longues distances. Du fait de leurs transmissions verticale comme horizontale ainsi que de la forte mobilité des Hommes ces interactions peuvent jouer des rôles considérables dans l'établissement des niches écoculturelles.

1.6.3 Répartition géographique des groupes humains

La mobilité potentiellement accrue des groupes humains nous conduit à mener les ECNM sur de vastes étendues géographiques. Ainsi biaisés par un Z' (voir figure 1.13(b)) pouvant recouvrir la quasi-totalité de l'environnement disponible (A), nous pourrions penser, à tort, que nous modélisons leur niche fondamentale, obtenant alors une niche modélisée potentielle (NMP) ne tenant pas compte ou peu des interactions biotiques. Or les groupes humains ont de vastes réseaux sociaux et leur présence en certains lieux donnés ne peut être possible ou envisagée sans interactions avec d'autres groupes humains. Les distributions de populations sont très influencées par ces relations amicales ou conflictuelles qui permettent la présence de ces dernières dans l'espace écologique Z'' dont l'échantillonnage influence l'issue de la modélisation.

Comme en témoigne la génétique humaine et la linguistique [Cavalli-Sforza *et al.*, 1994] la répartition des groupes humains actuelle est le fruit de plusieurs épisodes migratoires d'ampleurs et de longueurs différentes qui ont créé des mélanges biologiques et culturels expliquant la diversité actuelle comme les diversités du passé. D'autres méthodes procurent également des renseignements sur les migrations des groupes humains. On peut noter

ainsi, des études portant sur la variabilité des animaux ou des plantes domestiquées, sur la phylogénie des parasites ou des maladies humaines, comparant les teneurs isotopiques des ossements et portant sur du matériel archéologique lithique ou organique (voir [d'Errico *et al.*, 2012] pour une liste de références associées à ses diverses méthodes). Couplés à des techniques de datations, ces champs disciplinaires deviennent informatifs sur les directions et les durées des déplacements humains. Ces mouvements de population sont influencés par de nombreux facteurs comme le climat, la géographie, les ressources alimentaires, les relations biotiques (intra et inter-espèce) [d'Errico *et al.*, 2012]. Ces auteurs énoncent trois hypothèses pouvant expliquer les dispersions de groupes humains.

Hypothèse climatique bon nombre de distributions d'espèces est considéré comme résultant des changements environnementaux causés par des changements climatiques (ouverture de corridors, changement de niveau marins, barrière géographique, effet sur les ressources) qui réduisent ou facilitent l'accès à différentes régions. Les populations se trouvant isolées ou mélangées se distinguent dès lors des populations voisines. De plus sous des contraintes écologiques les populations peuvent être amenées à muter pour éviter de disparaître.

Hypothèse culturelle : des avancées technologiques (ou innovations) peuvent offrir la capacité à certains systèmes adaptatifs de s'affranchir des contraintes environnementales. Ces découvertes se produisent en des lieux et des temps variés et peuvent être transmises verticalement comme horizontalement. En ECNM l'hypothèse culturelle est généralement évoquée lorsque l'hypothèse climatique est rejetée.

Hypothèse sociale : des règles sociales peuvent jouer sur les comportements humains et ainsi avoir un impact sur les migrations humaines. Nous ne pouvons pas faire la distinction en ECNM entre hypothèse culturelle et hypothèse sociale.

Bien que le rôle de l'un de ces trois facteurs puisse être défini comme majoritaire dans l'établissement d'une population en un lieu donné, on ne peut pas exclure le fait que les deux autres facteurs aient pu jouer un rôle dans la distribution de cette population par le passé. Ainsi la répartition d'une culture peut, à première vue ne paraître dépendante seulement de facteurs sociaux-culturels mais résulter originellement de processus environnementaux.

1.6.4 Substrats environnementaux, expressions culturelles et cohésion sociale

Les groupes humains sont socialement fédérés par des liens fonctionnels et culturels. Ces groupes développent différentes stratégies pour survivre dans des milieux variés pouvant aller de la banquise aux déserts en passant par la forêt équatoriale. Les individus de ces groupes exploitent les ressources de lieux caractéristiques du mode de subsistance que le groupe a choisi de développer afin de survivre. Comme pour la biogéographie d'un taxon (définie selon son habitat et son rôle dans l'environnement), nous pouvons considérer l'étude des distributions géographiques de populations définies sur des traits fonctionnels

sociaux-économiques (chasseurs collecteurs, agriculteurs ou marchands nomades), également comme des niches écologiques, bien qu'il serait plus précis en ce cas de parler de niches « écofonctionnelles » car elles ne sont plus propres à l'espèce mais à des stratégies de subsistances caractéristiques d'une population de l'espèce. En ce sens, on constate que l'étendue géographique d'un métier particulier se trouve souvent corrélé à un environnement, comme des marchands nomades du désert, des pêcheurs de zones côtières ou des horticulteurs vivant en forêt équatoriale. Parmi ces populations exploitant différents environnements, s'immiscent des éléments fonctionnels spécifiques (matériels ou conceptuels) dans la vie quotidienne, dans le folklore, comme dans le verbal. Ces éléments introduits divergent en nature et en nombre selon les régions. Il en découle à long terme une diversité culturelle importante. À cela vient s'ajouter une variabilité au sein même de ces groupes qui peut avoir des origines fonctionnelles, culturelles, historiques ou induites par la proximité de groupes voisins.

Binford [2001], Diamond [1997], Wiens et Graham [2005] ainsi que Collard et Foley [2002] qui s'appuient sur l'ouvrage : *Atlas of World Cultures* de Price [1989] s'exercent à quantifier les variations géographiques de la diversité culturelle. Ces auteurs montrent que la densité des cultures humaines augmente exponentiellement avec la température, de façon linéaire avec les précipitations et va en diminuant de l'équateur aux pôles. Ils expliquent en partie cela par une accessibilité croissante aux ressources alimentaires due à l'augmentation de la biomasse à mesure que l'on s'approche de l'équateur. Ce deuxième point suggère une relation entre la surface de l'aire d'action du groupe et les conditions environnementales qui y règnent. En effet plus les ressources sont rares ou inatteignables, c'est-à-dire plus l'environnement est hostile, plus le groupe devra étendre son territoire afin de se les procurer en quantité suffisante pour survivre [Banks *et al.*, 2013b; Collard et Foley, 2002; Nettle, 1998]. Suivant ce raisonnement, Nettle [1998] souligne la nécessité de l'entraide (ou de la cohésion sociale) pour les groupes qui vivent en environnements défavorables. De cette manière, l'auteur met en avant l'importance de l'intercompréhension des groupes et établit ainsi une relation linéaire négative entre diversité linguistique et risque écologique. Notons par ailleurs que la diversité linguistique suit également le gradient latitudinal [Mace et Pagel, 1995]. Le long de ce gradient l'environnement évolue et l'accès aux ressources se trouve plus ou moins facilité. Des recherches basées sur les systèmes actuels de milieux ouverts occupés par des hommes montrent que la densité des populations humaines préhistoriques évoluait en fonction des climats, des zones accessibles et de la biomasse primaire qui constitue la nourriture des ongulés (principale ressource carnée de ces populations) [Delpech, 1999]. L'augmentation de biomasse reflète la diminution du risque écologique qui laisse libre cours aux dérives culturelles.

1.6.5 Un rythme d'évolution rapide

Contrairement à l'écologie des espèces animales, relativement stable sur de longues périodes car liée à leur évolution biologique, les systèmes culturels et fonctionnels humains

ont évolué de façon rapide et irrégulière au cours du temps [d'Errico *et al.*, 2012]. Ces changements comportementaux sont accompagnés d'innovations technologiques ou culturelles. Ces caractères sont héréditaires mais leur transmission et leur maintien varient en fonction de plusieurs facteurs : la démographie, l'efficacité des innovations, les contraintes environnementales et la perméabilité culturelle [Richerson et Boyd, 2005; Richerson *et al.*, 2009]. Des études corrélaient des innovations majeures comme l'apparition de la modernité culturelle [Ziegler *et al.*, 2013] à des changements climatiques rapides. À l'instar d'une mutation génétique, les innovations se propagent, mutent, se fixent ou disparaissent par sélection ou par dérive. Il en résulte une diversité culturelle importante entre populations qui se distinguent par leur technologies, leurs règles sociales (dont la complexité est variable), leurs valeurs symboliques et religieuses, leurs langues ainsi que par leurs identités ethniques [Banks *et al.*, 2008c]. Banks *et al.* [2013b] montrent que dans le cadre d'études confrontant populations néandertaliennes et populations d'hommes anatomiquement modernes, ces hypothèses fournissent une explication qui permet d'expliquer la diversité culturelle de façon plus appropriée et plus cohérente que des études comparatives au niveau biologique ou cognitif.

Les variations climatiques ont une fréquence relativement élevée notamment depuis 800 ka¹⁶ [Barker *et al.*, 2011 cités par Ziegler *et al.*, 2013]. Depuis, la lignée humaine a fortement évolué biologiquement et de nombreux faciès culturels ont vu le jour. Tester les relations entre manifestations de ces derniers et variations climatiques est une des possibilités qu'offre l'ECNM.

1.6.6 Portées et limites de l'ECNM

Modéliser les niches écoculturelles est instructif sur plusieurs aspects. Cette méthode permet d'évaluer dans quelles proportions l'apparition, le maintien ou la disparition d'une culture se trouve liée à l'environnement ou à des processus culturels. L'ECNM permet aussi d'observer l'évolution des niches au cours du temps.

1.6.6.1 Importance des variables environnementales

Une fois l'ECN d'une culture établie, la comparaison des résultats obtenus avec la distribution réelle de cette dernière peut nous indiquer si des facteurs environnementaux influencent la géographie de la culture. Dans le cas où ces facteurs sont relativement influents, il est alors possible de les classer selon leur importance, de les quantifier et de déterminer pour chacun quelles sont la ou les valeurs optimales ainsi que les valeurs extrêmes au-delà desquelles les conditions ne sont plus favorables à la culture.

16. 1 ka équivaut à un millier d'années

1.6.6.2 La part du culturel

Il arrive que les ECN prédites ne concordent pas avec la distribution réelle de la culture étudiée. Dans ce cas deux explications sont envisageables :

- la niche écoculturelle est bel et bien dictée par des phénomènes environnementaux, cependant nous ne les avons pas inclus dans la modélisation ou notre méthode ne permet pas de retracer leur relations ;
- la géographie culturelle est principalement influencée par des phénomènes sociaux ou culturels.

Généralement, car il est difficile de les évaluer, les relations biotiques ne sont pas incluses dans l'ECNM. Cependant, si suffisamment de facteurs abiotiques sont utilisés, l'on peut envisager que les combinaisons de certains d'entre eux arrivent à mimer les niches écologiques des organismes qui interagiraient avec la culture. De ce fait, il est possible de considérer qu'une part des relations biotiques serait indirectement modélisée (hypothèse du bruit étonnant [Araújo *et al.*, 2014; Soberón et Nakamura, 2009]). Les interactions avec d'autres cultures peuvent être considérées comme des relations biotiques. En ce sens, si les cultures qui ont un impact sur la niche de la culture étudiée sont régies principalement par des conditions environnementales, ces interactions seront alors prises en compte par la modélisation. Si tel n'est pas le cas et que l'influence de ces facteurs biotiques est conséquente, alors nous ne serons pas en mesure de modéliser la niche écoculturelle de la culture ciblée et concluons que les relations culturelles ont une part plus importante que l'environnement dans l'établissement de cette culture. Une situation intermédiaire peut être envisagée dans laquelle la projection géographique de la niche écoculturelle est bien plus étendue que l'aire d'influence étimée de la culture. Cette troisième situation, où la culture n'occupe qu'une partie de sa niche, peut être expliquée par la présence de barrières environnementales qui dressent les frontières du système adaptatif et des barrières culturelles au sein même de la niche.

Nous venons de voir que la modélisation de niche d'un trait culturel ne subissant aucune d'influence de l'environnement non culturel ne donne pas de résultat en accord avec la répartition connue de ce dernier. Dans ce cas, l'échec prédictif de l'ECNM signale l'importance des interactions culturelles. Le marqueur culturel (ou l'ensemble des marqueurs) utilisé pour l'analyse d'une telle population peut cependant avoir été à l'origine lié à l'environnement. Considérons par exemple, comme marqueur culturel la confection d'un objet. Cet objet qui aurait pu, au cours du temps subir des mutations morphologiques, usuelles ou fonctionnelles. L'objet peut donc se trouver, à un instant donné, détourné de son usage à tel point qu'il ne présente plus que très peu de similitudes avec l'objet originel, et donc ne plus être lié à l'environnement comme ce dernier avait pu l'être avant ces modifications. L'ECNM peut permettre la mise en évidence d'un tel phénomène en reconstruisant l'ECN pour ce marqueur culturel à une période précédente sous l'influence d'autres facteurs environnementaux.

1.6.6.3 Conservation et décalage de niches écoculturelles

Une fois l'ECN d'un système adaptatif reconstruite pour un environnement donné, il est possible de la projeter dans un autre environnement. La capacité de la niche projetée à prédire les données d'occurrences du second environnement nous informe quant à la conservation de ses caractéristiques ou d'éventuelles modifications (décalages, expansions, réductions).

Un système adaptatif peut être retenu, bien qu'inadapté, lorsqu'il se retrouve de façon subite dans un nouvel environnement (par exemple un changement climatique soudain), et ce même au détriment de la *fitness* des populations [Henrich, 2004].

Les apparitions d'innovations ou les influences d'autres cultures modifiant le système adaptatif peuvent être comparées à des mutations génétiques codantes. Nous pouvons ainsi continuer cette comparaison et mettre en évidence des phénomènes de fixations et de délétions.

1.6.6.4 Délétion ou fixation d'un trait culturel

L'ECNM permet de tester, pour une population donnée, si des événements environnementaux comme des changements climatiques favorisent l'apparition d'innovations (technologiques, symboliques, conceptuelles) ou l'émergence de modifications comportementales (changement d'économie, de mode de subsistance). Seulement quelques biais subsistent.

Premièrement la résolution de définition d'une culture ou d'un technocomplexe est variable en archéologie. Effectivement, le degré de détermination de ces derniers peut être défini, par exemple, selon quelques outils (caractéristiques : fossiles directeurs, ou confectionnés suivant une chaîne opératoire particulière : Discoïde, Levallois, ect) au Moustérien, selon des façonnages spécifiques de divers support osseux ou lithiques au Paléolithique supérieur, ou encore selon une production microlithique et des décors de céramiques au Néolithique. Ces différents niveaux de résolution peuvent constituer un obstacle aux comparaisons croisées des niches écoculturelles. Par ailleurs cette identification peut être faussée par des biais taphonomiques et épistémologiques.

Deuxièmement, l'héritabilité des caractères culturels est variable selon les relations du groupe à l'environnement [Banks *et al.*, 2008c]. De plus des innovations culturelles peuvent être adoptées mais, par la suite rapidement abandonnées, laissées de côté par effet de dérive, contre-sélectionnées, ou supplantées par d'autres traits culturels provenant d'une autre population.

Des traits culturels ne s'étant pas manifestés de façon matérielle se sont néanmoins fixés dans les populations. Certains sont décelables archéologiquement par le biais d'études biomoléculaires (génomés, protéomes), biochimiques (isotopes) ou des systèmes socio-économiques (exploitation des ressources, échanges, conflits). Tandis que d'autres traits

culturels immatériels peuvent être détectés par des études linguistiques (vocabulaire, expressions, grammaire) ou théologiques (mythologies, religions, légendes, superstitions).

Ainsi, des résidus de phénomènes culturels du passé peuvent être retrouvés dans des populations actuelles, subactuelles (documentées) ou archéologiques. Il est donc possible de reconstruire les niches de ces marqueurs culturels indirects afin d'apprécier leur relations éventuelles avec l'environnement et de reconstituer une part de l'histoire des populations qui les partagent.

Bien que l'ensemble des caractères culturels des populations du passé ne soient pas observable, l'ECNM, alliée à d'autres disciplines, procure un moyen d'évaluer l'étendue de l'ethnicité de populations même en l'absence de preuves matérielles. Toutefois les modélisations les plus robustes sont celles réalisées à partir de relevés d'occurrences de marqueurs culturels à la fois directs (éléments matériels) et indirects (gènes, langues, etc).

Dans ce travail, pour mener nos études (ECNM), nous employons des algorithmes prédictifs conçus pour les ENM. Dans la section suivante, nous faisons un état de l'art de l'emploi de ces derniers pour des études relevant de différentes disciplines (écologie, paléontologie, archéologie) avant de décrire plus en détail leur utilisation et leur fonctionnement dans le chapitre 2.

1.7 Exemples d'applications d'algorithmes prédictifs

Stimulées par la prise de conscience du changement climatique global, depuis la fin des années 90, les modélisations de niches écologiques se multiplient. Plus récemment, des études en paléoécologie et préhistoire ont vu le jour grâce à la publication de simulations de paléoclimats.

1.7.1 Utilisation des algorithmes prédictifs en écologie et paléoécologie

Les algorithmes prédictifs sont des outils permettant de prédire la présence d'espèces animales ou végétales en fonction de paramètres environnementaux. Leur utilisation pour la reconstruction de niches écologiques d'espèces protégées permet de mettre en place des plans de sauvegarde, de gestion, de conservation du milieu [Araùjo *et al.*, 2004; Loisele *et al.*, 2003; Ortega-Huerta et Peterson, 2004; Rushton *et al.*, 2004; Sánchez-Cordero *et al.*, 2005], ou de protection d'espèces menacées [Engler *et al.*, 2004; Norris, 2004]. Ces outils peuvent également servir à évaluer les conséquences d'un changement climatique pour une espèce donnée [Beaumont *et al.*, 2005; Bomhard *et al.*, 2005; Skov et Svenning, 2004; Thuiller *et al.*, 2005a,b], à prévenir l'expansion d'une maladie [Peterson *et al.*, 2004, 2014; Samy *et al.*, 2014] ou de l'invasion d'une espèce exotique [Peterson et Robins, 2003; Peterson et Vieglaes, 2001; Rose et Todd, 2014; Rouget *et al.*, 2004; Sandel et Dangremond, 2012], ou à observer des phénomènes évolutifs comme le décalage de niche [Peterson, 2011; Peterson et Li, 2014].

Grâce aux simulations climatiques et aux occurrences paléontologiques, les algorithmes prédictifs ont permis également de modéliser les niches écologiques d'espèces actuelles à des périodes reculées et de reconstituer les niches d'espèces disparues de nos jours. Les modèles ainsi établis permettent de connaître non seulement les répartitions potentielles mais aussi l'écologie de ces espèces. Des études ont été menées à différentes époques selon la disponibilité des simulations climatiques par exemple au Dernier Maximum glaciaire [Banks *et al.*, 2008a; Marske *et al.*, 2011; Nogués-Bravo, 2009; Svenning *et al.*, 2011], au Dévonien [Rode et Lieberman, 2005] ou encore à l'Ordovicien [Brame et Stigall, 2013; Stigall, 2012]. Enfin d'autres études diachroniques ont été conduites dans l'objectif de valider des hypothèses de conservations de niches comme celle menée par Lim *et al.* [2011] à propos des causes historiques responsables de la phylogéographie des oiseaux de Malaisie.

1.7.2 Utilisation des algorithmes prédictifs pour des groupes humains

La modélisation de niches écoculturelles (ECNM, voir section 1.6), a été évoquée la première fois par Banks *et al.* [2006]. Elle permet de mettre en évidence les éventuelles relations entre traits culturels et environnements puis, les cas échéant, de mettre en évidence les facteurs environnementaux les plus influents. L'ECNM permet également d'apprécier l'évolution des niches au cours du temps. Les traits culturels modélisés peuvent être de différentes natures et appartenir à des sociétés d'un passé plus ou moins lointain comme à une société actuelle ou sub-actuelle.

1.7.2.1 Reconstructions de niches écoculturelles préhistoriques

Banks *et al.* [2006] détaillent quelques exemples d'applications de l'ECNM dont certaines ont fait, par la suite, l'objet d'études plus approfondies [Anderson *et al.*, 2005; Banks *et al.*, 2008c; Gillam et Tabarev, 2006]. Gillam *et al.* [2007] déclarent que la modélisation de niches est un outil particulièrement adapté pour étudier la géographie des populations archéologiques car à l'instar des données biologiques, issues de collections muséales, de prospections ou de la littérature, les restes fossiles connus à ce jour ne représentent qu'une infime partie de ce qui a pu être.

Des études préliminaires, menées dans l'objectif de mieux comprendre les distributions des cultures du Paléolithique supérieur du vieux continent et celles des Paléoindiens d'Amérique du Nord [Anderson *et al.*, 2005; Banks *et al.*, 2006; Gillam et Tabarev, 2006] permettent d'émettre des hypothèses quant aux lieux d'Asie où les ancêtres des Paléoindiens pouvaient potentiellement vivre. Ces études exploratoires suggèrent, selon le principe de la conservation de niche (voir section 1.5), que l'origine des populations d'Amérique du Nord se situe, plutôt dans les régions australes d'Asie de l'Est que dans les régions septentrionales sibériennes. Suite à ce peuplement succéderait, 1000 à 2000 années après, une vague migratoire (dite post-Clovis) en Alaska et au Canada provenant de Sibérie ayant pu subvenir par la voie continentale (modèle du collier de perle [Anderson et Gillam, 2000]) ou par cabotage

[Birdsell, 1951 cité par Gillam *et al.*, 2007]. Dans ces études les différents groupes culturels ont été définis selon des critères de technologie lithique : pointes cannelées, pointes Folsom ou pointes de Cumberland [Anderson *et al.*, 2005; Gillam et Tabarev, 2006]. Des études postérieures utilisent également de tels artefacts comme marqueurs culturels [Banks *et al.*, 2008c, 2009], ou s'appuient sur des assemblages comprenant plusieurs types de vestiges archéologiques, afin de rendre plus pertinentes les appartenances des groupes étudiés aux technocomplexes [Banks *et al.*, 2013a, 2011, 2013b].

d'Errico et Sánchez Goñi [2003] puis Sepulchre *et al.* [2007] mettent en relation la disparition des Néandertaliens et les variations climatiques abruptes. Partant de ce résultat Banks *et al.* [2008b] reconstruisent les ECN des Néandertaliens ainsi que celles des premiers Hommes anatomiquement modernes d'Europe. Ces travaux permirent d'évaluer la largeur de leurs niches respectives au cours de trois périodes différentes : préH4, H4 et G18¹⁷. Les prédictions et les projections des niches des homininés dans les trois environnements respectifs à ces périodes permettent aux auteurs de conclure, sous l'hypothèse de conservation de niches, que l'effet compétiteur des populations aurignaciennes est un des facteurs principaux causant la disparition des populations néandertaliennes.

Banks *et al.* [2008c] se sont ensuite penchés sur l'ECNM de deux technocomplexes du Paléolithique supérieur. Les auteurs ont ainsi trouvé que les niches écoculturelles des populations solutréennes sont plus froides et plus humides que celles occupées par les populations épigravétiennes. L'étude permet aussi d'observer que ces dernières faisaient face à un risque écologique moindre comparé à celui que subissaient les populations solutréennes. Ainsi, cette étude suggère que des variations adaptatives peuvent coïncider avec des variations culturelles.

Partant de l'observation d'une non-circulation de matières premières entre deux populations de culture badgoulienne¹⁸, Banks *et al.* [2011] appliquent l'ECNM afin de voir si les deux populations partagent ou non la même niche. Les résultats montrent qu'il n'y a pas de différence significative entre les deux niches écoculturelles modélisées, mais qu'une nuance est cependant notable. Devant les deux territoires définis, les auteurs émettent alors l'hypothèse d'un décalage de niche se mettant en place afin de réduire la compétition entre ces populations. Ils soulignent également que ce phénomène pourrait expliquer la régionalisation des populations mise en place lors du Paléolithique supérieur.

Deux ans plus tard, [Banks *et al.*, 2013b] utilisent l'ECNM afin de voir si les variations de cultures matérielles coïncident avec les variations climatiques et si, dans un tel cas, des variations de niches écologiques peuvent être observées. Pour cela, les auteurs reconstruisent les niches écoculturelles de deux technocomplexes : Protoaurignacien et Aurignacien ancien (déterminables par leurs industries lithiques respectives et cloisonnées dans des périodes pouvant être identifiées par datations au carbone 14 suivies d'un traitement bayésien). Les

17. Périodes délimitées par des événements géo-climatiques dits de Heinrich correspondant à des variations climatiques causées par des débâcles d'icebergs.

18. Le Badgoulien est une culture préhistorique du Paléolithique supérieur de l'Ouest de l'Europe.

auteurs remarquent alors un élargissement de niche. Cette transformation semblant être concomitante à la détérioration climatique du HS4, ils estiment alors que les différences relevées au niveau des cultures matérielles sont des réponses adaptatives aux changements climatiques. Au vu de ce résultat ils soulignent l'intérêt d'étudier les niches ECN de divers technocomplexes de transitions.

[Banks *et al.* \[2013a\]](#) mettent en évidence les relations entre deux mouvements culturels du néolithique et l'environnement. Cette étude est décrite chapitre [10](#).

Enfin, [d'Errico et Banks \[2013\]](#) montrent que l'étude des trajectoires culturelles de populations humaines à l'échelle régionale permettrait d'identifier les évolutions de niches écoculturelles.

1.7.2.2 Reconstructions de niches écoculturelles historiques et de sociétés traditionnelles actuelles

Dans ce document nous détaillons des ECNM effectuées sur des populations actuelles (chapitre [8](#)) et historiques (chapitre [9](#))

Synthèse du chapitre : Théories et concepts écologiques

Le concept de la niche « hutchinsonnienne » se complexifie largement lorsqu'on intègre les contributions biologiques comme l'exclusion compétitive [Gause, 1934] ou les interactions positives comme la facilitation ou le mutualisme [Callaway *et al.*, 2002; Travis *et al.*, 2005; Yamamura *et al.*, 2004]. La niche écologique possède sa propre dynamique qui évolue au gré des changements environnementaux d'ordre biotique ou abiotique. Dans ce contexte, la conceptualisation de l'espace écologique représentée par le diagramme BAM permet de mieux cerner l'enveloppe écologique modélisée à partir des données d'occurrences et environnementales dont on dispose.

Dans ce chapitre, nous avons vu que la modélisation de niches écologiques, qui incorpore les facteurs environnementaux, permet un apport d'informations beaucoup plus varié et précis que la distribution géographique basée sur un simple relevé de présence ou d'absence. Nous nous sommes interrogés sur les notions ambiguës de découpe de l'espace écologique et avons retenu la compartimentation par niche comme étant la plus adéquate à la modélisation avec, notamment, la notion d'hypervolume permettant de conceptualiser les niches comme des interactions de facteurs environnementaux mesurables et comparables.

Nous avons décrit différents biais pouvant découler de l'échantillonnage des occurrences et dressé une liste des espaces écologiques modélisés en fonction de la représentativité de l'écologie par les occurrences.

Enfin, nous avons décrit le concept de niche écoculturelle (ECN) puis montré que la part de l'environnement et du culturel dans la mise en place, le maintien ou l'abandon d'un système adaptatif pouvait être apprécié grâce à la modélisation de niche écoculturelle (ECNM [Banks *et al.*, 2006]).

Les algorithmes prédictifs de niches écologiques

Préambule

Dans l'objectif d'initier les lecteurs non familiers aux techniques de modélisation de niches, ce chapitre décrit l'utilisation et le fonctionnement des algorithmes prédictifs. Pour cela, il s'appuie fortement sur l'ouvrage de [Franklin \[2009\]](#) qui offre une synthèse précise des différentes techniques de modélisation ainsi que sur le livre de [Peterson et al. \[2011\]](#). Enfin, après une revue des différentes méthodes pour évaluer la pertinence des prédictions obtenues, nous verrons comment optimiser ces dernières.

Sommaire

2.1	Modélisation de niches écologiques à l'aide d'algorithmes prédictifs . . .	48
2.2	Les Algorithmes prédictifs en détails	48
2.3	Représentativité des données	59
2.4	Traitement des données avant la modélisation	68
2.5	Optimiser une prédiction	70
2.6	Stochasticité, validité et pertinence des prédictions	73

Comme dans le chapitre précédant, nous nous situerons au rang de l'espèce afin d'expliquer notre raisonnement. Tous les individus constituant notre population seront donc des occurrences de l'espèce. Dans l'objectif de rendre plus concrets la méthodologie et les divers points techniques abordés dans les sections de ce chapitre, nos propos seront illustrés dès que possible par un exemple récurrent concernant la modélisation de niches écologiques d'un rongeur arctique, le *Dicrostonyx* sp.¹. Chaque paragraphe faisant référence

1. L'animal est plus communément connu sous le nom de lemming car confondu avec un autre rongeur, le *Lemmus* sp.)

à cet exemple sera précédé du symbole ✂. Cette étude a permis également d'appliquer et de valider la méthodologie d'ENM proposée dans ce document sur une espèce animale, connue pour être un bon marqueur climatique [Hinton, 1926; Kowalski, 1995; Prost *et al.*, 2010]

2.1 Modélisation de niches écologiques à l'aide d'algorithmes prédictifs

Une façon de déterminer la répartition géographique d'un groupe d'individus est de recenser tous les lieux où ces derniers sont présents. Ceci est envisageable pour une espèce actuelle mais impossible à réaliser en archéologie. Beaucoup d'auteurs listent des occurrences d'espèces ou créent des bases de données pour dresser ensuite des atlas. Mais aussi complets que leurs relevés puissent être, ils ne sont jamais exhaustifs et par conséquent les aires de répartitions décrites ne peuvent être considérées comme exactes. Les approches phylogénétiques et linguistiques se heurtent également à la difficulté d'obtenir un échantillon représentatif. Ainsi, les résultats obtenus par ces approches ne peuvent être que descriptifs. La modélisation permet de surmonter cela. En effet, si l'on arrive à caractériser quels sont les critères d'établissement d'un groupe dans certaines régions, alors on peut extrapoler l'ensemble des autres régions. Une telle modélisation peut être effectuée grâce à l'utilisation d'algorithmes prédictifs de niches écologiques.

2.2 Les Algorithmes prédictifs en détails

Dans cette section après avoir vu leur fonctionnement général nous passerons en revue différentes familles d'algorithmes prédictifs et détaillerons le fonctionnement et l'emploi de deux d'entre eux fréquemment utilisés : MaxEnt et GARP. Nous évoquerons aussi *Open Modeller* qui présente l'intérêt d'intégrer plusieurs de ces algorithmes et qui offre la possibilité d'automatiser des tâches fréquemment effectuées en ENM. Un emploi de cette fonctionnalité est présenté en détail dans le chapitre 5.

✂ Nous nous sommes servi d'*Open Modeller* afin de reconstruire les niches écologiques de plusieurs espèces de rongeurs pour la période actuelle et pour la période correspondant au Dernier Maximum glaciaire (DMG). Ces modélisations ont été effectuées avec la collaboration d'Aurélien Royer qui a fourni les données d'occurrences.

2.2.1 Fonctionnement général des algorithmes prédictifs

Il existe deux approches pour reconstruire des niches écologiques : l'approche mécaniste et l'approche corrélative (section 1.4.2). Nous nous intéresserons seulement à la deuxième car le but de cette thèse est d'utiliser la modélisation de niches pour des populations du

passé. Si nous voulions mener une approche mécaniste, il faudrait pouvoir simuler ces populations et apprécier leurs variations lorsqu'on modifie les variables environnementales, chose actuellement impossible en l'état de nos connaissances. En revanche les vestiges archéologiques couplés aux simulations paléoclimatiques constituent des jeux de données suffisants pour envisager des modélisations par approche corrélative.

Les modélisations de niches écologiques par approche corrélative (figure 1.12) établissent des corrélations entre les occurrences d'une espèce en un lieu donné et son environnement. Les outils permettant cela se nomment algorithmes prédictifs et requièrent deux types de données que nous nommons entrées, par opposition à la sortie qui est la niche écologique prédite.

2.2.1.1 Première entrée : les occurrences

Il s'agit de l'information signalant la présence ou l'absence de la population² étudiée. Elle se présente sous la forme d'une liste de n coordonnées géographiques indiquant chacune, pour un lieu donné, la présence ou l'absence de l'espèce. Il existe deux types d'algorithmes, ceux qui requièrent des données de présence et d'absence et ceux qui ne nécessitent que des données de présence. Les seconds sont qualifiés d'algorithmes *presence only*³. Notons que l'évaluation de ces différents types d'algorithmes par Brotons *et al.* [2004] préconise l'emploi des données d'absence pour obtenir de meilleurs résultats lorsque ceci est possible. Notre travail étant à vocation archéologique, nous focaliserons notre attention sur les algorithmes prédictifs pouvant fonctionner seulement avec des données de présence. Effectivement, dans la discipline, les données d'absence ne peuvent pas être prises en compte car l'absence de traces d'un phénomène ne signifie pas qu'il ait été absent. De plus, les données de présence sont plus fiables que les données d'absence qui peuvent être dues à divers biais comme un effort d'échantillonnage insuffisant ou un manque d'information (figure 2.1).

✍ Les données d'occurrences fossiles de *Dicrostonyx* sp. ($n = 46$) proviennent de la thèse de Royer [2013]. Les données d'occurrences actuelles du rongeur ($n = 204$) ont été collectées dans la base de données GBIF (www.gbif.org [Flemons *et al.*, 2007]). Il s'agit d'occurrences de présence seulement.

2.2.1.2 Deuxième entrée : les couches environnementales

Les données environnementales qui constituent la deuxième entrée de la modélisation de niche doivent être sous forme de couches géoréférencées. Elles sont nommées à ce propos couches environnementales. Ces données correspondent à une variable environnementale (précipitations, températures, altitude, etc). Ces couches sont de type *raster*, cartes

2. Le terme de population est à considérer dans son sens statistique. En archéologie par exemple il pourrait s'agir de vestiges diagnostiques d'une culture donnée.

3. présence seulement

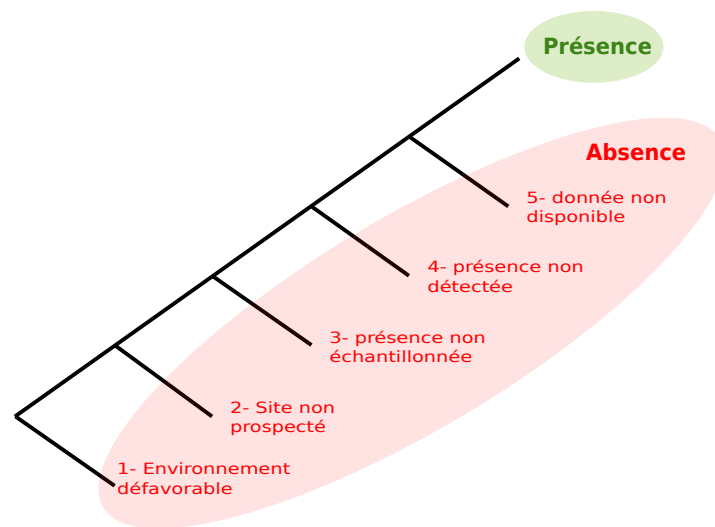


FIGURE 2.1 : Arbre de probabilité des données de présence, d'après un cours en ligne [Peterson, 2013] (<http://biodiversity-informatics-training.org>, modifié). Dans cette figure sont représentés cinq cas d'absence dont quatre d'entre eux ne sont pas liés aux conditions environnementales. L'auteur montre aussi que les données de présence sont beaucoup plus informatives que les données d'absence.

géographiques quadrillées qui possèdent une valeur de la variable qu'elles représentent en chaque point de grille. Le maillage (ou résolution) le plus grossier d'entre elles détermine la résolution (ou la précision) de la niche obtenue en sortie. En ENM, les données environnementales utilisées le plus communément sont de nature abiotique.

✍ 26 couches environnementales ont été utilisées pour notre cas école : 19 climatiques (Worldclim, [Hijmans *et al.*, 2005]) et 7 topographiques dérivées de l'élévation (ETOPO1 [Amante et Eakins, 2009])

Remarque. Dans la suite du document, nous adopterons le vocabulaire suivant :

- Concernant les sorties des algorithmes, nous parlerons de prédictions ou de projections.
- Les couches environnementales (ou enveloppes, ou grilles, ou *rasters*) sont composées de carrés que l'on dénommera aussi points de grille ou pixels.
- Les termes résolution, maillage, précision ou grain seront employés à même escient.
- Une présence ou un point désignera une occurrence de présence de l'espèce.
- De la même manière que dans le chapitre 1, nous expliquerons la méthodologie employée en prenant pour exemple la modélisation de niche écologique au niveau de l'espèce.

Le schéma figure 2.2 représente le fonctionnement général des algorithmes prédictifs. Initialement prévus pour des modélisations de niches écologiques lorsqu'ils utilisent des données d'occurrences d'espèce et des données environnementales d'ordres climatiques et topographiques, ces algorithmes permettent, en changeant le type d'entrées, de reconstruire

Tableau 2.1 : Exemples de types de niches pouvant être modélisés par les algorithmes prédictifs. Les méthodes de modélisation peuvent être utilisées à différentes fins. Pour l’instant seules les niches basées sur des données écologiques (en l’espèce climatiques et topographiques) ont été réalisées (ligne 1 à 4) les autres lignes du tableau sont des études qui pourraient être envisagées sous réserve de leur cohérence.

	Données environnementales	Données d'occurrences	Type de modélisation	Références
1	Écologiques	Taxon actuel	Niche écologique	ex : [Sandel et Dangremond, 2012]
2	Écologiques	Taxon fossile	Niche paléoécologique	ex : [Svenning <i>et al.</i> , 2011]
3	Écologiques	Culture	Niche écoculturelle	ex : [Banks <i>et al.</i> , 2011]
4	Écologiques	Groupe linguistique	Niche écolinguistique	[Antunes <i>et al.</i> , 2013]
5	Écologiques	Trait fonctionnel	Niche écofonctionnelle	
6	Écologiques	Représentation	Niche écosymbolique	
6	Biologiques	Culture	Niche bioculturelle	
7	Biologiques	Groupe ethnolinguistique	Niche bioethnolinguistique	
8	Biologiques	Trait fonctionnel	Niche biofonctionnelle	
9	Économiques	Culture	Niche économicoculturelle	
10	Économiques	Groupe ethnolinguistique	Niche économicoethnolinguistique	
11	Économiques	Trait fonctionnel	Niche économicofonctionnelle	
12	Économiques	Politiques	Niche politicoéconomique	
13	Économiques	Sociologiques	Niche socioéconomique	

des niches à dimensions autres qu’écologiques (tableau 2.1). Par exemple, des niches à dimension culturelle [Antunes *et al.*, 2014; Banks *et al.*, 2013a, 2011, 2006, 2008b,c, 2009; d’Errico et Banks, 2013] ou linguistique comme abordé dans le chapitre 8. Le tableau 2.1 montre que d’autres utilisations des algorithmes prédictifs de niches écologiques sont concevables à condition de disposer des données requises.

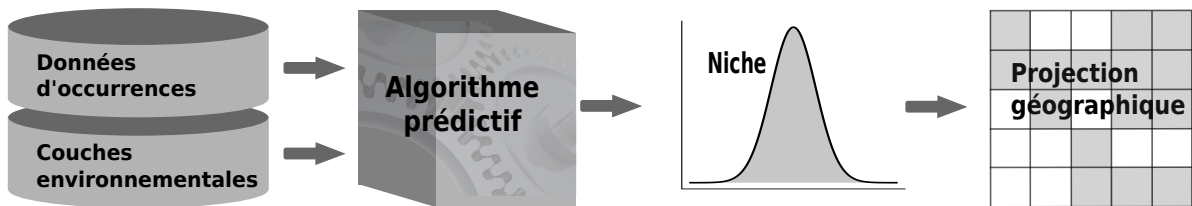



FIGURE 2.2 : Fonctionnement général des algorithmes prédictifs. Tous les algorithmes prédictifs abordés dans ce document de thèse fonctionnent en prenant deux types d’entrées : les occurrences géographiques (c’est-à-dire les coordonnées géographiques) et les couches environnementales (incluant une donnée numérique en chaque point de grille). Les niches modélisées en sorties seront ensuite projetées géographiquement sous forme de grille de données numériques géoréférencées de résolution au moins égale à la moins précise des enveloppes environnementales.

Tableau 2.2 : Prédications et projections de deux niches dans deux environnements. Ce tableau résume les différentes possibilités de prédictions et de projections pour deux niches modélisées dans deux environnements. Le nombre de représentations spatiales est de quatre : deux prédictions et deux projections.

Niche prédite		Environnement	
		CTRL	LGM
Niche prédite	Actuelle	Prédiction de <i>Dicrostonyx</i> sp. CTRL	Projection de <i>Dicrostonyx</i> sp. CTRL dans LGM
	Fossile	Projection de <i>Dicrostonyx</i> sp. LGM dans CTRL	Prédiction de <i>Dicrostonyx</i> sp. LGM

2.2.2 Prédications et projections de niches

Une fois la niche modélisée, c'est-à-dire une fois que l'algorithme prédictif a établi la formule de l'hypervolume de la niche (section 1.2.3.2), l'algorithme fournit une représentation cartographique en appliquant la formule à chaque pixel d'une carte. On parle de projection ou d'expression géographique de la niche. À ce stade, les valeurs environnementales injectées dans la formule de la niche peuvent être celles de l'environnement dans lequel vit l'espèce et à partir duquel la niche a été calculée, ou elles peuvent correspondre à un environnement d'une autre région ou d'une autre époque (passée comme future). Par convention les termes « prédiction » désigne l'expression géographique de la niche calculée dans son propre environnement et « projections » les expressions géographiques de cette niche dans d'autres environnements (voir tableau 2.2 pour un exemple .

La figure 2.3 montre comment à partir d'un calcul de niche on peut obtenir plusieurs projections. Nous remarquons dans cette figure qu'une projection peut avoir lieu sur des environnements dont l'étendue varie ; par exemple, l'environnement 3 est beaucoup plus étendu que les deux autres. Cette propriété peut permettre de faire des estimations d'aire de répartition potentielle de l'espèce pour des environnements ayant une résolution différente de celle de la couche environnementale utilisée pour la modélisation.

En gestion et en conservation de la biodiversité, les projections de niches sont souvent utilisées afin d'estimer la répartition potentielle d'une espèce invasive dans un nouveau milieu ou dans un nouvel espace accessible [Peterson *et al.*, 2008b] mais aussi afin d'évaluer les chances de succès d'une réintroduction d'espèce [Martínez-Meyer *et al.*, 2006].

2.2.3 Liste non exhaustive des algorithmes prédictifs, descriptions et spécificités

Il y a plusieurs catégories d'algorithmes prédictifs de niches, chacun procède différemment : intelligence artificielle, régressions statistiques, arbres de décisions, analyses multifactorielles, etc. Certains prennent en compte toutes les données pour établir un mo-

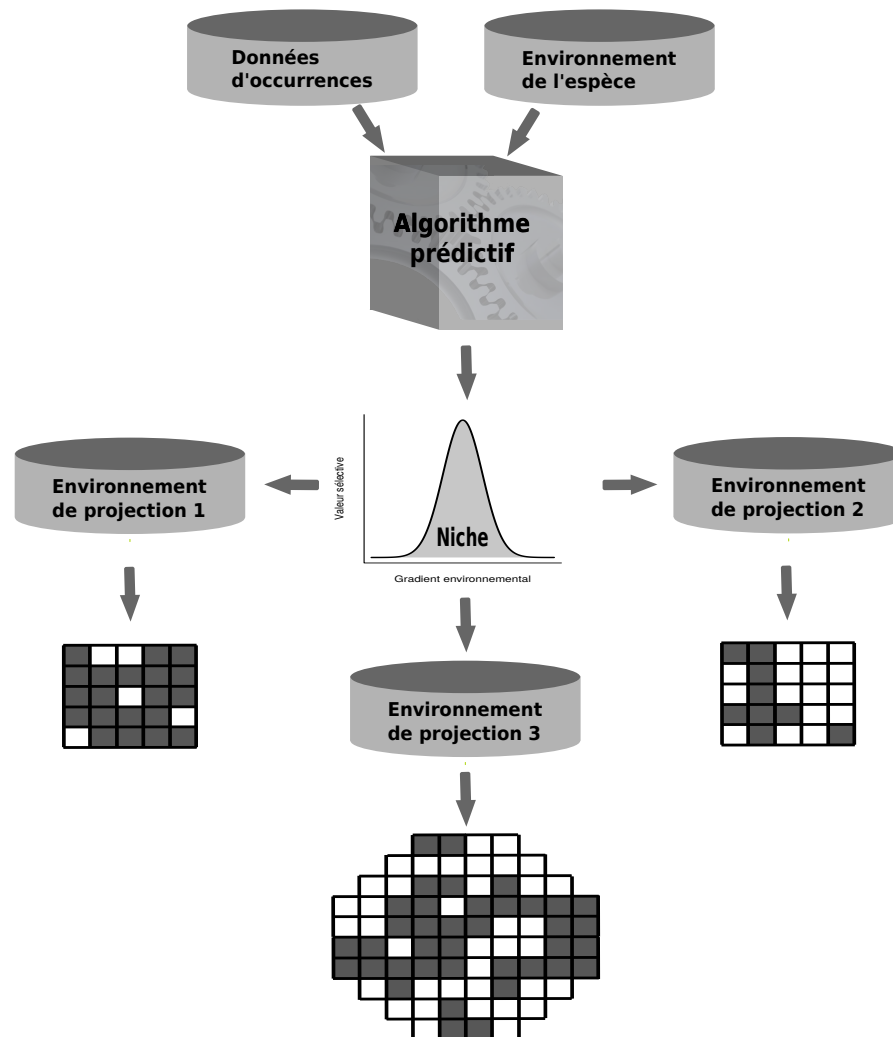


FIGURE 2.3 : Projections de niches écologiques. L'algorithme prédictif établit la niche dans un premier temps en fonction des occurrences et de l'environnement d'origine de l'espèce. Puis une il la projette dans d'autres environnements. Ici l'environnement 1 et 2 sont deux environnements concernant la même région mais à des périodes différentes et l'environnement 3 est celui d'une autre région.

dèle alors que d'autres procèdent par apprentissage supervisé⁴. Le tableau 2.3 dresse une liste non exhaustive d'algorithmes prédictifs dont l'usage est le plus répandu. La plupart d'entre eux sont disponibles indépendamment dans des logiciels de traitements statistiques, mais il existe également des logiciels qui proposent une utilisation simultanée (tableau 2.4). Dans le cadre de cette thèse nous avons utilisé *Open Modeller* qui permet de lancer ces algorithmes en lignes de commande (voir chapitre 5).

2.2.3.1 Régressions statistiques

Les méthodes les plus intuitives et les premières utilisées pour trouver les relations entre des occurrences et l'environnement furent les régressions linéaires comme les GLM⁵ [McCullagh et Nelder, 1989], ou MARS⁶ [Friedman, 1991]. Bien que ces méthodes donnent des résultats relativement corrects, elles ne sont pas toujours en adéquation avec l'écologie car souvent les systèmes suivent des règles complexes et multifactorielles. C'est pourquoi l'emploi de méthodes ayant la possibilité d'intégrer des régressions non linéaires comme les GAM⁷ [Hastie et Tibshirani, 1990] ont été rapidement privilégiées [Thuiller, 2003].

2.2.3.2 Arbres de décisions

La modélisation par arbres de décisions, en particulier de type CART [Breiman *et al.*, 1984], permet à la suite d'un apprentissage automatique une classification hiérarchique des facteurs environnementaux selon leur importance dans l'établissement de la niche écologique.

2.2.3.3 Analyses multifactorielles

Ce type d'analyse (ACP⁸, AFC⁹, etc.) a été souvent employé car il permet des regroupements, cependant il ne devient pertinent que lorsque les données ont une grande fiabilité et en grand effectif.

4. Méthode qui consiste à scinder en deux lots les relevés de présence-absence, un lot utilisé pour l'apprentissage (*training*) alors que l'autre est réservé à l'application (*testing*). La proportion de ces deux lots est en général déterminée par l'utilisateur, pour ce faire ce dernier tient compte, de l'effectif critique relatif à l'algorithme utilisé (voir l'article de [Hernandez *et al.*, 2006]).

5. *General linear models*

6. *Multivariate Adaptive Regression Splines*

7. *General additive models*

8. Analyse en Composantes Principales

9. Analyse Factorielle des Correspondances

Tableau 2.3 : Liste non exhaustive de méthodes utilisées pour la prédiction de niches écologiques.

Algorithme	Nom/description	Exemples d'applications	Références statistiques
ANN	<i>Artificial Neural Networks</i>	[Lek <i>et al.</i> , 1996; Manel <i>et al.</i> , 1999]	[Hopfield, 1982; McCulloch et Pitts, 1943]
Bioclim		[Ward, 2007]	[Beaumont <i>et al.</i> , 2005; Nix, 1986]
BM	<i>Bayesian Model</i>	[Pereira et Itami, 1991]	[Strahler, 1980]
ANUCLIM	<i>Bioclimatic Profiles</i>	[Elith et Burgman, 2002]	[Houlder <i>et al.</i> , 2000]
BRT	<i>Boosted Regression Trees</i>	[Leathwick <i>et al.</i> , 2006]	[Elith <i>et al.</i> , 2008]
CART	<i>Classification and Regression Tree analysis</i>	[Dettmers <i>et al.</i> , 2002; Fertig <i>et al.</i> , 2002; Franklin, 1998; Thuiller, 2003; Vayssières <i>et al.</i> , 2000; Walker, 1990]	[Breiman <i>et al.</i> , 1984]
CCA	<i>Canonical Correspondence Analysis</i>	[Guisan <i>et al.</i> , 1999]	[Braak, 1988]
CTA	<i>Classification Tree Analysis</i>	[Thuiller <i>et al.</i> , 2003]	[Breiman <i>et al.</i> , 1984]
DA	<i>Discriminant Analysis</i>	[Dettmers <i>et al.</i> , 2002; Manel <i>et al.</i> , 1999; Mastrotillo <i>et al.</i> , 1997; Sattler <i>et al.</i> , 2007]	[Zingg, 1990]
Domain		[Ward, 2007]	[Carpenter <i>et al.</i> , 1993]
ENFA	<i>Ecological Niche Factor Analysis</i>	[Sattler <i>et al.</i> , 2007; Václavík et Meentemeyer, 2009]	[Hirzel <i>et al.</i> , 2002]
ENVDIST	<i>Environmental Distance</i>	[Chen, 2008]	[Farber et Kadmon, 2003]
ENVSORE	<i>Envelope Score</i>	[Silva <i>et al.</i> , 2014]	[Nix, 1986]
FEM	<i>Fuzzy Envelope Model</i>	[Robertson <i>et al.</i> , 2004]	[Robertson <i>et al.</i> , 2004]
GAM	<i>Generalized Additive Models</i>	[Bio <i>et al.</i> , 1998; Elith et Burgman, 2002; Franklin, 1998; Pearce et Ferrier, 2000; Thuiller, 2003; Yee et Mitchell, 1991]	[Hastie et Tibshirani, 1990]
GARP	<i>Genetic Algorithm for Rule-set Production</i>	[Elith et Burgman, 2002; Peterson <i>et al.</i> , 2002]	[Stockwell, 1999a,b; Stockwell et Noble, 1992]
GBM	<i>General Boosting Method</i>	[Leathwick <i>et al.</i> , 2008]	[Ridgeway, 2006]
GLM	<i>Generalized Linear Models</i>	[Bio <i>et al.</i> , 1998; Guisan <i>et al.</i> , 1999; Lek <i>et al.</i> , 1996]	[McCullagh et Nelder, 1989]
GRASP	<i>Generalized Regression Analysis and Spatial Prediction</i>	[Lehmann <i>et al.</i> , 2002]	[Lehmann <i>et al.</i> , 2002]
Logistic R	<i>Logistic Regression</i>	[Gehring et Swihart, 2003]	[Hosmer et Lemeshow, 1989]
LSR	<i>Least Square Regression</i>	[Thuiller <i>et al.</i> , 2006]	[Draper et Smith, 1981]
MARS	<i>Multivariate Adaptive Regression Splines</i>	[Mateo <i>et al.</i> , 2010]	[Friedman, 1991]
MaxEnt	<i>Maximum Entropy</i>	[Raxworthy <i>et al.</i> , 2007]	[Guisu et Shenitzer, 1985; Jaynes, 1957; Phillips <i>et al.</i> , 2004]
MD	<i>Mahalanobis distance method</i>	[Dettmers <i>et al.</i> , 2002]	
MDA	<i>Mixture Discriminant Analysis</i>	[Broennimann <i>et al.</i> , 2007]	[Hastie et Tibshirani, 1996]
N-Bayes	<i>Naive Bayes</i>	[Guo et Liu, 2010]	[Duda <i>et al.</i> , 2001]
PCA-based	<i>Principal Component Analysis</i>	[De Meyer <i>et al.</i> , 2008]	[Robertson <i>et al.</i> , 2001]
RF	<i>Random Forests</i>	[Evans et Cushman, 2009]	[Breiman, 2001; Cutler <i>et al.</i> , 2007]
RS	<i>Rought set</i>	[Guo et Liu, 2010]	[Pawlak, 1991]
SI	<i>Surface Interpolators</i>	[Pereira et Itami, 1991]	[Pereira et Itami, 1991]
SVM	<i>Support Vector Machine</i>	[Drake <i>et al.</i> , 2006]	[Burges, 1998; Schölkopf <i>et al.</i> , 2001, 2000; Vapnik, 1995]

2.2.3.4 Intelligence artificielle

Les algorithmes prédictifs de niches écologiques actuels intègrent des techniques relevant de l'intelligence artificielle comme des algorithmes génétiques ou des réseaux de neurones [McCulloch et Pitts, 1943]. Ils permettent de retrouver des règles complexes entre l'environnement et un faible jeu d'occurrences.

2.2.4 Caractère stochastique des prédictions

Les spécificités, la nature stochastique et l'emploi des divers algorithmes produisent des modèles hétérogènes. En effet, bien que modélisées à partir des mêmes entrées, deux niches produites par deux algorithmes différents ne sont pas identiques. Les niches produites sont uniques et peuvent ainsi varier de façon importante selon les algorithmes employés, mais également selon les paramétrages de ces derniers. Au fil des études, on découvre ainsi des spécificités et/ou des tendances propres aux algorithmes. Par exemple, il a été constaté la tendance de certains à concentrer les zones de prédictions de niches autour des occurrences, on parle dans ce cas d'algorithmes « microprédictifs » (exemple MaxEnt) alors que d'autres attribuent une aire plus grande à la niche, dans ce cas ils sont qualifiés, comme GARP, d'algorithmes « macroprédictifs » [Peterson *et al.*, 2008a].

Ne connaissant pas au préalable quel sera l'algorithme fournissant le résultat le plus en accord avec la réalité, les modélisateurs ont pris l'habitude de croiser pour chaque étude les prédictions de niches issues de plusieurs d'entre eux. Certains logiciels comme *Open Modeler* intègrent plusieurs algorithmes et offrent la possibilité d'en incorporer de nouveaux, de plus ces interfaces (graphiques ou en ligne de commande) permettent d'automatiser ces tâches. Toutefois parmi la vingtaine de techniques existantes, deux d'entre elles sont plus fréquemment utilisées.

2.2.5 Les algorithmes préférentiellement utilisés

MaxEnt et GARP sont deux algorithmes fréquemment utilisés pour la modélisation de niches écologiques lorsque l'on dispose uniquement de points de présence. Leur utilisation est répandue notamment pour deux raisons : la première est qu'ils fournissent des modèles de niches pertinents et la deuxième est qu'ils furent les premiers à pouvoir être exécutés via une interface graphique relativement simple d'utilisation (MaxEnt en 2004 est doté d'une interface java [Phillips *et al.*, 2004]¹⁰, et GARP possédant sa propre interface graphique DesktopGARP depuis 1999 [Stockwell, 1999b]¹¹). Les deux algorithmes procèdent de façon itérative. Lors de chaque itération, ils partitionnent les données en jeux d'apprentissages et jeux de tests. Nous dénommons ces itérations : *runs*. Les utilisateurs disposent de jeux d'op-

10. Java Runtime Environment qui a l'avantage d'être portable (multi plate-forme)

11. interface non portable

tions propres à chaque algorithme, choix de règles, de seuils ou de méthodes statistiques pouvant conduire à l'obtention de modèles très différents selon les paramétrages.

L'utilisation des deux algorithmes est quasiment équivalente : une interface graphique permet à l'utilisateur d'entrer ses données, de choisir des paramètres (tels le nombre d'itérations ou des seuils de sélections) et fournit en sortie des fichiers directement utilisables en SIG. Les deux algorithmes offrent la possibilité d'effectuer un *Jackknifing* sur les variables environnementales (voir section 2.4.2.1). Cette opération consiste à générer des modèles en écartant les variables une à une. Ce qui permet d'apprécier leur poids dans l'élaboration des modèles et de voir s'exprimer les variables secondairement importantes.

La durée d'une expérience est dépendante du nombre de *runs*, des paramètres sélectionnés, des seuils déterminés, de la taille de l'échantillon et enfin de la puissance de calculs de l'ordinateur utilisé. Maxent est souvent utilisé car il a été démontré comme performant pour divers *scenarii* de modélisation en écologie, biogéographie ou conservation [Elith et Leathwick, 2009]

2.2.5.1 MaxEnt

MaxEnt [Elith *et al.*, 2011; Phillips *et al.*, 2006, 2004] est la contraction des mots *Maximum* et *Entropy*. Cet algorithme cherche à définir une probabilité d'occurrences pour l'espèce étudiée en chaque pixel de la région d'étude en partant du principe que la solution la plus optimale expliquant un phénomène inconnu est celle qui maximise l'entropie. C'est-à-dire celle qui fournit la distribution des probabilités de présence la plus parcimonieuse en fonction des contraintes imposées. [Peterson *et al.*, 2008a]. En ENM, ces contraintes consistent à maintenir le plus bas possible les différences entre les valeurs moyennes des variables environnementales correspondantes aux pixels prédits et les valeurs moyennes observées pour les occurrences.

2.2.5.2 GARP

L'algorithme GARP a été développé par Holland [1975] puis implémenté dans l'interface graphique DesktopGARP pour la modélisation de niche écologique [Stockwell, 1999a,b; Stockwell et Noble, 1992]. GARP est l'acronyme de *Genetic Algorithm for Rule set Prediction*. Comme indiqué, il s'agit d'un algorithme génétique, c'est-à-dire qui s'inspire du concept de l'évolution par sélection naturelle, où les solutions apparaissent de manière aléatoire parmi de nombreuses itérations puis sont par la suite retenues. L'algorithme crée des jeux de règles environnementales aléatoirement. Ces règles sont assimilées à des gènes. Lors des premiers *runs*, des gènes représentant des solutions plus ou moins pertinentes sont générés. L'algorithme affecte à ces gènes des mutations et les recombine entre eux. Les nouveaux gènes obtenus sont alors comparés aux premiers et seuls les plus pertinents sont retenus. Ainsi, au fur à mesure la pertinence des prédictions augmente jusqu'à qu'elles atteignent un degré de fiabilité jugé optimal.

Tableau 2.4 : Liste de logiciels dédiés à l'application des algorithmes prédictifs (liste non exhaustive).

Logiciels	Algorithmes inclus dans le logiciel	Références
<i>Open Modeller</i>	Bioclim, GARP, ENFA, ENVSCORE, ENVDIST, MaxEnt, ANN, RF, SVM, (autres*)	[De Souza Muñoz <i>et al.</i> , 2011]
Biomod2 (<i>package R</i>)	GLM, GBM, GAM, CTA, ANN, SRE, FDA, MARS, RF, MaxEnt	[Thuiller <i>et al.</i> , 2013]
Dismo (<i>package R</i>)	MaxEnt, Bioclim, Domain, MD	[Hijmans <i>et al.</i> , 2013]
Ensemble		
ModEco	GLM, CART, ANN, MaxEnt, SVM, Domain, N-Bayes, Bioclim, Rought set	[Guo et Liu, 2010]
DesktopGARP	GARP	[Stockwell, 1999a]

(*)les algorithmes cités sont livrés avec le logiciel *Open Modeller* mais cette architecture offre aussi la possibilité d'implanter d'autres algorithmes.

L'atout majeur des algorithmes à intelligence artificielle est qu'ils observent une large gamme des solutions possibles et convergent vers la plus optimale. Ils évitent ainsi, grâce à de multiples itérations, de lisser (moyenner) les prédictions selon les tendances régionales ou de rester dans un optimum local.

2.2.5.3 Logiciels permettant l'exécution de plusieurs algorithmes

Il existe plusieurs logiciels qui permettent l'exécution d'algorithmes prédictifs (tableau 2.4), parmi cette liste *Open Modeller* et *Biomod* semblent offrir le plus de possibilités.

Open Modeller, [De Souza Muñoz *et al.*, 2011], a été spécialement conçu pour l'emploi simultané de plusieurs algorithmes. L'utilisateur de ce logiciel a la possibilité de sauvegarder ces paramétrages pour de futures analyses. Le logiciel est fourni avec plusieurs algorithmes déjà inclus, dont GARP et MaxEnt, et offre la possibilité d'en rajouter (voir tableau 2.4). Le mode «om_console» permet une exécution en ligne de commande. Cette fonctionnalité permet une automatisation des tâches. Il est alors possible d'effectuer une multitude de prédictions sans commettre d'erreurs, ce mode permet aussi l'emploi de statistiques à la volée comme le *partial ROC*, l'AUC, l'*accuracy* (voir chapitre 5 pour voir comment nous avons procédé à l'automatisation des tâches pour les études présentées dans cette thèse). Enfin *Open Modeller* offre la possibilité de créer des consensus des prédictions obtenues. Cependant le nombre de prédictions doit être limité à cinq et ces dernières doivent être au préalable rendues binaires¹². Nous verrons chapitre 6, comment nous avons créé des consensus sans ces limitations.

Biomod [Thuiller *et al.*, 2009] est un *package R* qui permet l'exécution de différents algorithmes et l'évaluation de leur performance selon plusieurs statistiques (comme : TSS, ROC, AUC ou Kappa). Ce *package* récemment mis à jour (*Biomod2*, [Thuiller *et al.*, 2013]) offre de nombreuses possibilités pour l'automatisation de modélisations, la création de consensus, l'évaluation de modèles et la comparaison de prédictions.

12. probabilité nulle ou maximale, une prédiction de la sorte ne possède plus de nuance.

2.2.6 Les algorithmes les plus adaptés aux études archéologiques

Il est possible de déterminer la niche écologique d'une espèce de différentes manières. Alors que des relevés sur le terrain pour une espèce actuelle permettraient d'avoir des données de type présence/absence, nous ne disposons, en archéologie, que des traces témoignant d'une présence. Les méthodes de modélisation dont des données de type présence/absence sont requises ne sont donc pas adaptées à cette discipline. On privilégie donc les modélisations dites *presence only* (qui ne nécessitent pas de données concernant l'absence d'une espèce) ou générant eux-mêmes statistiquement de « fausses » absences. Ces algorithmes sont, dans ce cas, dits à pseudo-absence. Bien que l'on dispose avec ces méthodes de moins d'informations, on évite par ailleurs, les erreurs d'observations qui consisteraient à déclarer une absence en un point où il y a pu avoir une présence qui n'a pas laissé d'empreinte archéologique. En effet, on ne peut pas affirmer qu'il n'y a pas eu présence de telle ou telle culture à un endroit donné par le simple fait de sa non-observation, l'absence de preuve n'étant pas une preuve d'absence.

Une fois la méthode de modélisation appropriée définie, nous devons porter un regard sur la représentativité des données.

2.3 Représentativité des données

Une prédiction de niche : sortie de la modélisation, est d'autant plus fiable que la qualité des entrées est haute. Pour obtenir des résultats optimaux, il est donc important de choisir les points d'occurrences comme les variables environnementales avec objectivité et minutie.

2.3.1 Qualité des points d'occurrences

Comme vu dans la section 2.2.1, les points d'occurrences constituent une des deux entrées principales pour la modélisation de niche. La qualité du modèle va donc de pair avec la qualité de ces données. Il est donc important de s'interroger avant toute modélisation, sur la représentativité de l'échantillon. La qualité d'un tel échantillon pour la prédiction d'une niche écoculturelle, constitué de relevés de présence uniquement, dépend de paramètres liés à l'effectif et à l'identification de la population étudiée.

2.3.1.1 Taille de l'échantillon

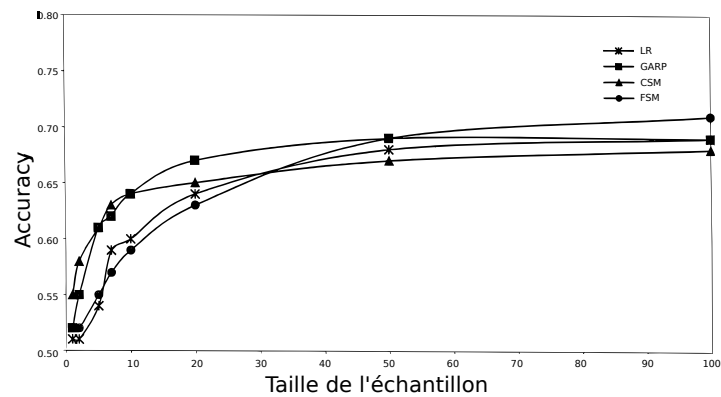
L'effectif des relevés de présence est déterminant pour deux raisons : d'une part, ce nombre doit être suffisamment important afin de rendre légitime la constitution d'un groupe (quantité suffisante d'individus possédant un ou plusieurs caractères communs) ; d'autre part, comme les algorithmes nécessitent un minimum d'occurrences pour produire des résultats robustes (figure 2.4), la taille de l'échantillon doit atteindre un minimum requis spécifique à l'algorithme employé.

La taille de l'échantillon nécessaire pour la représentativité d'un groupe varie selon la rareté et la variabilité intergroupe du ou des caractères considérés. L'effectif minimum est donc à évaluer au cas par cas et est souvent remis au jugement d'un spécialiste en la matière. En revanche, l'effectif doit toujours être supérieur au minimum requis par les algorithmes prédictifs utilisés [Hernandez *et al.*, 2006; Stockwell et Peterson, 2002] de sorte à s'assurer que la pertinence des modèles obtenus ne puisse être remise en question. Dans les faits, le nombre d'occurrences n'est pas une limitation forte car d'après les travaux de Stockwell et Peterson ainsi que de Hernandez *et al.*, nous voyons, figure 2.4, que certains algorithmes sont capables de générer des prédictions fiables à partir de dix occurrences et deviennent très pertinentes au-delà de 25. Ces auteurs ont montré pour 103 espèces d'oiseaux (dans l'étude de Stockwell et Peterson) et pour 18 espèces appartenant à des classes différentes (dans l'étude de Hernandez *et al.*), que la taille de l'échantillon ne nécessitait pas d'être colossale afin d'obtenir des prédictions fiables. Ces constats sont une manne pour des champs disciplinaires comme l'archéologie et la paléontologie, où les données sont souvent peu abondantes (voir section 2.3.1). Toutefois, la taille de l'échantillon critique peut être difficile à atteindre car en plus de leur rareté, ces occurrences doivent se trouver suffisamment espacées les unes des autres. Effectivement, si deux occurrences se situent dans le même point de grille alors, dans ce cas, une seule présence sera comptabilisée en ce point de grille pour la modélisation. Il est donc nécessaire de retirer ces relevés avant de modéliser une niche, d'autant plus que la présence de multiples occurrences en un même pixel peut fausser les statistiques permettant l'évaluation des prédictions (pour plus de détails sur ces statistiques voir la section 2.6).

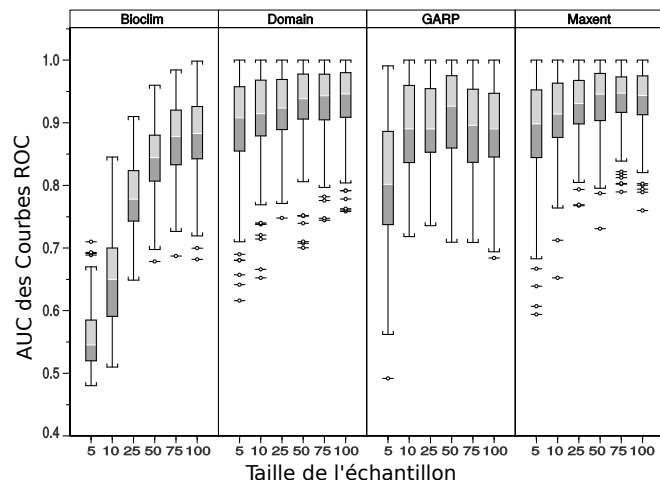
✎ Pour la modélisation de la niche du *Dicrostonyx* sp. actuel, la totalité des occurrences relevées était de 303. Beaucoup d'entre elles, proches les unes des autres, se trouvaient dans le même point de grille. Après traitement, c'est-à-dire après avoir retiré les occurrences spatialement redondantes, nous comptons encore 207 occurrences spatialement uniques, effectif largement suffisant pour écarter l'hypothèse d'un biais induit par la taille de l'échantillon.

2.3.1.2 Précision des attributions culturelles

Au-delà du problème de l'effectif, peut subvenir un biais d'échantillonnage causé par une détermination erronée de l'entité que l'on cherche à modéliser. À l'inverse de l'identification d'espèces qui s'appuie sur des clés de détermination précises et péremptaires, la catégorisation de groupes humains en fonction de traits culturels est moins triviale. En effet, les critères de détermination d'une espèce biologique sont généralement immuables, car génétiques, alors que des identificateurs culturels peuvent être empruntés ou encore détournés de leur fonction primaire par certains groupes (section 1.6). Ces caractères, marqueurs culturels, peuvent également être absents dans d'autres groupes appartenant à la même population car ils ont été substitués par un ou plusieurs éléments provenant d'une autre population ou bien n'ont pas été encore adoptés. Ce phénomène est d'autant plus



(a) Évaluation de la performance des algorithmes : LR, GARP, FSM et CSM selon leur *accuracy* (d'après [Stockwell et Peterson, 2002]) ; plus l'*accuracy* est proche de 1 plus la prédiction de l'algorithme est considérée comme fiable.



(b) Évaluation de la performance des algorithmes : Bioclim, Domain, GARP et MaxEnt selon leur l'AUC (aire en dessous de la courbe ROC) d'après [Hernandez *et al.*, 2006] ; plus l'AUC est proche de 1 plus la prédiction de l'algorithme est considérée comme fiable.

FIGURE 2.4 : Effet de la taille de l'échantillon sur les prédictions de niches écologiques.

marqué pour les populations archéologiques où ces marqueurs culturels ne peuvent être que d'ordre matériel et donc sur lesquels s'ajoutent des incertitudes quant à leur fonction. À cela s'additionne une incertitude concernant l'absence de pièces déterminantes, comme un fossile directeur, pouvant être d'origine taphonomique ou due à un l'effort d'excavation insuffisant. Pour éviter les erreurs d'identifications induites par les biais précédemment cités les archéologues se basent sur des technocomplexes, c'est-à-dire des ensembles d'éléments dont la présence conjuguée de plusieurs d'entre eux certifie une affectation culturelle¹³ à un mode de vie ou une culture particulière. Ainsi, dans de nombreux cas les interprétations ne peuvent dépasser le stade d'hypothèses et les sites ne peuvent pas être considérés comme appartenant au groupe culturel et par conséquent inclus dans les modélisations. D'autre biais réduisent également la taille de l'échantillon comme l'effort de prospections, la superficie du territoire couvert, l'incertitude des datations ainsi que les dogmes établis épistémologiquement.

2.3.1.3 Données parasites, bruit et taphonomie

À l'inverse des autres disciplines, les données archéologiques et paléontologiques peuvent être biaisées pour différentes raisons. Par exemple, une espèce présente dans un niveau archéologique ne signifie pas forcément que l'espèce étaient présente en ce lieux à l'époque correspondante à ce niveau, la strate ayant pu subir des modifications géologiques ayant pu provoquer un mélange du matériel qu'elle contenait avec celui d'une couche adjacente correspondant à une autre époque durant laquelle l'espèce était, cette fois, présente. Il en est de même pour les marqueurs culturels. Ces remaniements de la stratigraphie (il peut s'agir également de palimpseste¹⁴) peuvent être le fruit de plusieurs facteurs : des perturbations géologiques (par exemple cryoturbations, solifluxions, glissements, effondrements, drainages), des bioturbations (animaux fouisseurs et racines), comme des perturbations anthropiques (réutilisation d'outils, habitats troglodytes, creusement de fosses sépulcrales, fouilles clandestines).

Dans d'autre cas, une espèce ou un marqueur culturel peut disparaître, ou sa détection peut être rendue impossible à cause de l'altération des vestiges. Ainsi dans ces disciplines étudiant des populations passées, les absences ne constituent pas une information fiable. Dans la figure 2.5, on note au moins cinq exemples de relevés d'absence erronées possibles. Cette figure montre aussi que les présences sont également moins fiables que s'il s'agissait de populations actuelles car des processus post-dépositionnels peuvent induire des présences dans des lieux inopportuns. Les archéologues et paléontologues doivent alors redoubler

13. Pour attribuer une affiliation culturelle à un site il n'est pas nécessaire qu'il ait délivré l'ensemble des éléments caractéristiques du technocomplexe. En revanche, la présence d'un seul ou de très peu d'objets de cette liste, aussi représentatifs soient-ils, ne prouve pas avec certitude l'appartenance au groupe culturel suspecté.

14. Pièces archéologiques de différentes époques se trouvant à un même niveau suite à une lacune sédimentaire, des épisodes d'érosion ou d'autres événements géologiques.

de vigilance et ne considérer uniquement que les présences dont la véracité a été vérifiée par plusieurs méthodes comme : des datations, des études techno-typologiques des objets archéologiques associés, des comparaisons d'altération de surfaces, ainsi que des études taphonomiques et de formation des sites.

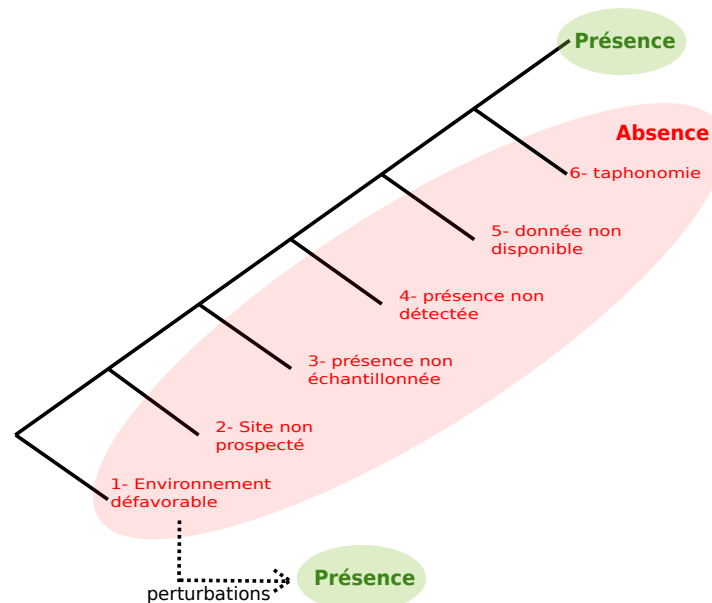


FIGURE 2.5 : Arbre de probabilité des données de présence et d'absence dans un concept archéologique ou paléontologique. La taphonomie des vestiges (6) peut masquer des données de présence et les perturbations stratigraphiques des sites des perturbations peuvent quant à elles fournir de fausses présences (c'est-à-dire des présences où règnent des conditions environnementales inappropriées).

2.3.2 Qualité des données topographiques

Parmi les entrées, les données topographiques sont les plus précises et les plus simples à acquérir. En effet, pour l'actuel nous disposons de nombreux moyens de mesures terrestres, aériens ou satellitaires pour les observer. À partir des relevés d'altitudes, par exemple ETOPO [Amante et Eakins, 2009], découle le calcul de plusieurs variables topographiques (tableau 2.5). Ces variables sont calculées en fonction de la différence d'élévation d'une cellule et de ces huit cellules environnantes.

Les variables topographiques pour les périodes actuelles sont donc très fiables et très précises. Pour les périodes anciennes, au regard des échelles de modélisations (les plus fines

Tableau 2.5 : Liste de variables topographiques. Ces variables s’obtiennent à partir d’un modèle numérique de terrain, pour cela nous avons utilisé la fonction *terrain()* (du *package Raster* de R) qui permet le calcul de ces variables selon la méthode décrite par [Wilson *et al.*, 2007]. Ces variables sont nommées comme ceci dans toutes les modélisations effectuées dans cette thèse.

Code	Description	unité
dem	<i>Digital elevation model</i> . Modèle numérique de terrain, il correspond à l’Élévation (ou altitude) moyenne d’un pixel	m
tpi	<i>Topographic Position Index</i> . L’indice de topographie est la différence entre la valeur d’une cellule et la moyenne des valeurs des huit cellules qui l’entourent. Cette valeur renseigne sur la position topographique d’un pixel (par exemple font d’une vallée, milieu d’une pente ou sommet d’une crête) et peut fournir des informations sur la géomorphologie d’une région	-
tri	<i>Terrain Ruggedness Index</i> . L’index topographique de rugosité est la moyenne des valeurs absolue des différence entre une cellule et les huit cellules qui l’entourent. Cette valeur est basse dans les régions planes et haute dans les autre cas	-
rough	Rugosité du paysage. Différence entre la valeur minimale et la valeur maximale des huit cellules environnantes d’un pixel.	-
flowdir	Direction du ruissellement	Degrés ou radians
slope	Pente	Degrés ou radians
aspect	Orientation de la pente ou assiette (si la pente est nulle alors l’aspect est affecté de la valeur $\frac{\pi}{2}$ ou 90°)	Degrés ou radians

étant de 30" ¹⁵) et compte-tenu de la lenteur des processus géologiques ¹⁶, nous partons du principe que la morphologie du paysage est sensiblement équivalente à celle que l’on connaît de nos jours depuis. Toutefois, nous corrigeons l’altitude pour les périodes à fortes régressions marines puis nous recalculons les autres variables qui en dépendent comme expliqué dans le tableau 2.5.

✍ Ainsi, dans notre exemple, pour estimer la topographie de zones actuellement sous l’eau au DMG, nous corrigeons l’altitude selon les travaux de [Siddall *et al.*, 2003]. À cette période le niveau marin est inférieur d’environ 120 m par rapport à nos jours. Nous soustrayons donc 120 mètres au modèle numérique de terrain actuel et calculons les autres variables topographiques comme indiqué tableau 2.5.

Remarquons que, si pour les régions continentales, du fait de leur lenteur, les processus tectoniques et l’érosion sont peu perceptibles aux échelles considérées pour les modéli-

15. La résolution s’exprime en degrés et est divisée en minutes puis en secondes, ainsi une résolution de 30 secondes correspond à des pixels d’environ 1 km² à l’équateur dont la superficie diminue à mesure que l’on s’approche des pôles.

16. Dans cette thèse nous ne reconstruisons pas de niches de populations vivant à des périodes antérieures à 21000 BP, les processus géologiques n’ont pas eu le temps d’induire des mutations observables lorsqu’on regarde le paysage avec une acuité de l’ordre du kilomètre.

sations de niches que nous effectuons, les zones côtières ont pu subir des érosions plus importantes par l'océan ou des comblements par apports fluviaux. Nous devons donc être vigilant quant aux interprétations des prédictions pour ces régions, notamment s'il s'avère que la topographie joue un rôle majeur dans la distribution des populations étudiées.

2.3.3 Qualité des données climatiques

En modélisation de niche, on emploie en général des données climatiques en plus des données topographiques, mais les résolutions de celles-ci influencent fortement la nature de la niche qui en résulte (voir chapitre 1). Pour des modélisations de niches actuelles, il est possible d'obtenir des données climatiques précises alors que pour des modélisations concernant d'autres périodes (futures ou passées) les climats doivent être modélisés.

2.3.3.1 Utilisation de modèles climatiques

Il existe différentes façons de modéliser le climat. Les reconstructions ou les réanalyses sont produites à partir de mesures et d'observations directes alors que les simulations proviennent de la modélisation du climat issue de l'interpolation de mesures indirectes de type *proxies*¹⁷, de transformations selon des principes géophysiques (volume des calottes glaciaires, circulation thermohaline, trajectoires orbitales et axiales du globe terrestre), ainsi que différents forçages (solaire, volcanique ou gaz à effet de serre). Les modèles climatiques peuvent être établis selon différents couplages :

- des circulations générales de l'atmosphère et des océans (par exemple IPSL CM4 et CM5A, [Dufresne *et al.*, 2013; Kageyama *et al.*, 2013]) ;
- de l'atmosphère et de la colonne d'eau des 50 premiers mètres de profondeur des océans [Kang *et al.*, 2008] ;
- de l'atmosphère et des températures de surfaces des océans [Kageyama *et al.*, 2005].

Bien que l'exactitude des simulations soit moins rigoureuse que celle des reconstructions ou des réanalyses, les simulations sont effectuées sur de longues périodes, fournissant une valeur de la variable étudiée avec une période variant de 6 à 24 heures (appelée « pas de temps ») en chaque point de grille. De telles simulations nécessitent l'emploi de supercalculateurs et souvent la résolution finale reste assez grossière compte tenu du temps de calcul. De plus les problématiques des climatologues sont généralement centrées sur les mécanismes généraux aux échelles régionales, voire mondiale, qui ne justifient pas l'emploi de données à très haute résolution spatiale. De ce fait, les simulations de climats possédant une résolution adaptée à la modélisation de niches et correspondant aux époques et aux lieux recherchés n'ont pas toujours été calculées.

17. Un *proxy* est un enregistrement local permettant d'évaluer le climat, par exemple une carotte glaciaire ayant piégé du dioxygène dont le rapport isotopique ($\delta^{18}\text{O}$) renseigne sur la température qui régnait lors du piégeage de ce gaz.

2.3.3.2 Résolution spatiale

Les couches environnementales doivent être des *rasters* (section 2.2.1) dont chaque point de grille contient une valeur. Le *raster* idéal, serait la moyenne sur plusieurs années des mesures effectuées pour le paramètre climatique en chaque point de grille, or les *rasters* atteignent vite un nombre important de cellules et une telle collecte d'informations s'avérerait être un travail titanesque. C'est pourquoi nous travaillons à partir de reconstructions, de réanalyses ou de simulations climatiques.

Le principe de fonctionnement des algorithmes prédictifs est dans un premier temps de modéliser la niche écologique (établir la formule de l'hypervolume) et dans un second temps de projeter géographiquement ce modèle (section 2.2.1), la résolution des dimensions environnementales détermine celle de la projection, en effet la résolution du modèle sera dans le meilleur des cas égale à la résolution de la couche environnementale de résolution la plus grossière. Contrairement à la topographie pour laquelle nous disposons de moyens permettant de représenter les données en couches dotées d'un maillage très fin, les grilles climatiques, à la résolution bien plus larges, limitent la précision des niches écologiques modélisées.

Néanmoins, le grain peut être affiné en utilisant des techniques d'interpolations ou de régressions basées sur des *proxies* ou des mécanismes climatiques comme le *downscaling*¹⁸ dynamique, le *downscaling* empirique, ou le *downscaling* statistique ([Levavasseur *et al.*, 2011; Vrac *et al.*, 2007], voir chapitre 4).

2.3.3.3 Résolution temporelle

Pour apprécier la dynamique des niches au cours du temps, l'idéal serait de posséder des données journalières continues d'une période A à une période B, or mis à part certaines reconstructions permettant de remonter jusqu'à la fin du XIX^e siècle (date des relevés quotidiens les plus anciens), nous ne possédons aucun enregistrement de la sorte couvrant des zones d'études suffisamment larges pour permettre la modélisation des populations nous intéressant. Des simulations mensuelles ou journalières du climat sur des périodes longues (par exemple : dernier millénaire) offrent la possibilité, quant à elles, d'obtenir un haut degré de finesse chronologique et permettent ainsi d'apprécier la variabilité climatique sur des périodes courtes (voir chapitre 4).

Afin de conduire une ENM ou une ECNM pour une population archéologique ou paléontologique, il est nécessaire de posséder des variables environnementales à une très haute résolution spatiale et correspondant à l'époque de la population ciblée, or de telles données ne sont pas toujours disponibles. Nous proposons dans cette thèse une méthode statistique, voir chapitre 4, qui permet d'obtenir pour une époque donnée une simulation à très haute résolution à partir d'une simulation possédant un maillage inadapté à la modélisation de niche.

18. Terme anglais pouvant être traduit par descente d'échelle.

Tableau 2.6 : Liste des variables bioclimatiques *WorldClim*, selon [Hijmans *et al.*, 2005]. Notons qu'il existe des études incorporant plus de variables bioclimatiques (ex : 35 variables pour Bioclim v5.1 [Beaumont *et al.*, 2005] qui comprend les 19 variables décrites dans cette table plus 16 autres en relation avec l'humidité et les radiations solaires.).

Code	Description
bio1	Température moyenne annuelle
bio2	Amplitude thermique quotidienne moyenne
bio3	Isothermalité (bio2 / bio7*100)
bio4	Saisonnalité de la température (coefficient de variation)
bio5	Températures maximales du mois le plus chaud
bio6	Températures minimales du mois le plus froid
bio7	Amplitude annuelle des températures (bio5 – bio6)
bio8	Températures moyennes du trimestre le plus humide
bio9	Températures moyennes du trimestre le plus sec
bio10	Températures moyennes du trimestre le plus chaud
bio11	Températures moyennes du trimestre le plus froid
bio12	Total des précipitations annuelles
bio13	Précipitations du mois le plus humide
bio14	Précipitations du mois le plus sec
bio15	Saisonnalité des précipitations (coefficient de variation)
bio16	Précipitations du trimestre le plus humide
bio17	Précipitations du trimestre le plus sec
bio18	Précipitations du trimestre le plus chaud
bio19	Précipitations du trimestre le plus froid

La nature des variables climatiques influence également la modélisation de niche. Pour conduire nos études nous optons pour l'utilisation de variables bioclimatiques.

2.3.3.4 Utilisation de variables bioclimatiques

Les variables bioclimatiques sont des variables calculées selon les températures, les précipitations et la saisonnalité, cette nature leur confère un sens biologique [Hijmans *et al.*, 2005] et sont par conséquent très souvent utilisées par les écologues ([Bernardo-Silva *et al.*, 2012; Lim *et al.*, 2011; Peterson *et al.*, 2007]). Le tableau 2.6 énumère la liste des variables de *WorldClim* [Hijmans *et al.*, 2005]. Ces variables globales peuvent être téléchargées (<http://www.worldclim.org/>) pour plusieurs périodes à différentes résolutions ou bien peuvent être construites à partir de climatologies (voir section 4.4.1.9).

✍ Pour la modélisation des niches écologiques des lemmings nous avons téléchargé sur *WorldClim* [Hijmans *et al.*, 2005] les variables bioclimatiques issues du *downscaling* du modèle CCSM4 [Gent *et al.*, 2011] à une résolution de 2.5'.

L'emploi des variables bioclimatiques est pertinent pour la modélisation de l'écologie des organismes [Nix, 1986], ainsi dans cette thèse, afin d'observer la dynamique des niches et l'impact des fluctuations climatiques sur ces dernières, nous nous efforçons d'obtenir des variables bioclimatiques avec une résolution spatiale et chronologique adéquate.

2.4 Traitement des données avant la modélisation

Il est parfois nécessaire de sélectionner les données avant la modélisation pour obtenir un résultat de meilleure qualité.

2.4.1 Traitement des occurrences

Certains algorithmes intègrent la notion de densité en accordant une importance supérieure aux occurrences de plus forte densité, dès lors que les occurrences leur aient été transmises avec une probabilité de présence équivalant à la densité de l'organisme en ce lieu. D'autres, n'intégrant pas cette fonctionnalité, mais du fait qu'ils procèdent par sous-échantillonnage, peuvent être biaisés par des présences multiples en un même point de grille. De plus, dans nos démarches nous cherchons à connaître l'écologie d'un organisme, or rien ne prouve qu'une forte densité de ce dernier en un lieu signifie qu'il lui est plus convenable qu'un autre. En effet, beaucoup d'espèces en condition de stress adoptent une stratégie de reproduction, ou se réfugient dans les lieux les plus proches où la cause du stress est absente. D'autre part, concernant des données du passé la taille des populations, quoi qu'elle puisse être modélisée (par exemple [Bocquet-Appel *et al.*, 2005; Hassan, 1981; Kline et Boyd, 2010]), reste difficile à apprécier. Afin d'éviter de biaiser la modélisation, on s'assure alors que les occurrences de présence ne représentent que la présence d'un seul organisme par point de grille. Ce traitement d'information est une fonctionnalité¹⁹ de base intégrée dans bon nombre de logiciels permettant l'exécution d'algorithmes prédictifs.

2.4.2 Traitement des variables environnementales

2.4.2.1 Sélection de variables

Dans l'objectif d'éviter le sur-apprentissage, il arrive que des modélisateurs de niches, s'assurent, avant d'effectuer une modélisation, que les variables environnementales ne soient pas fortement corrélées entre elles. Ces derniers préconisent cette opération notamment lorsqu'il s'agit de variables bioclimatiques car elles peuvent être nombreuses ($n=36$ [Nix, 1986] ou $n=19$ [Hijmans *et al.*, 2005]). Selon ces auteurs, cette sélection a un double effet sur les modèles de niches produits par les algorithmes prédictifs. En effet, cela évite à la fois l'inclusion de biais dans la modélisation et permet de surcroît une simplification de l'interprétation des résultats.

Généralement, ce traitement des variables environnementales est effectué grâce à des tests de Spearman qui permettent d'évaluer les corrélations entre elles. La méthode de Johnson *et al.* [2002] consiste à tester les corrélations par paire de variables, d'éliminer une variable par paire corrélée au-dessus d'un seuil (par exemple $r > 90\%$, [Araújo *et al.*, 2014]). À ce stade, le modélisateur choisi de ne conserver que la variable qui paraît expliquer le

19. Traitement des points pour qu'ils soient spatialement uniques (option : *spatially unique*).

mieux la biologie de l'espèce ou la plus facile à interpréter. Ces opérations sont renouvelées jusqu'à ce que les scores de corrélations de toutes les paires de variables restantes possibles soient inférieurs au seuil prédéfini [Rissler et Apodaca, 2007]. Il a été développé, en 2010 un *package* R : *polycor* [Fox, 2010]) qui permet d'établir plusieurs types de corrélations (corrélation de Spearman, polysérialles et polychoriques). Ce package a notamment été utilisé à cet effet par Norris [2014].

Pour les modélisations de niches menées dans cette thèse, nous avons choisi de ne pas effectuer cette sélection de variables environnementales pour plusieurs raisons :

- La corrélation n'est pas identique en tout point d'une carte car des variables non intégrées lors de la création du modèle pour des lieux donnés peuvent entrer en jeu dans d'autres lieux.
- Bien que statistiquement on puisse obtenir des différences faibles entre des variables, il faut noter que sur de grandes étendues (une zone d'étude pouvant atteindre plusieurs dizaines de milliers de pixels) les différences peuvent être accentuées (un grand écart-type) ou concentrées dans des régions de taille relativement importante de la zone d'étude.
- Un des intérêts de la modélisation de niche est d'identifier les facteurs environnementaux qui la composent, étant conscients des relations complexes pouvant exister, nous ne voulons pas incorporer de subjectivité dans la méthode que nous employons.
- La plupart des algorithmes que nous utilisons sont dotés de fonctions telles que la *Jackknifing* permettant d'écarter un ou plusieurs prédictors et de tester les différences entre les modèles obtenus avec un jeu de variables restreint ainsi que leur cohérence. De ce fait, peu d'algorithmes sont réellement susceptibles d'être affectés par des variables redondantes.
- Dans l'objectif de projeter les niches prédites pour une population dans des régions ou à des époques différentes (donc susceptibles de posséder des conditions climatiques différentes) il est nécessaire de travailler avec un jeu de variables le plus complet possible car les corrélations entre ces dernières peuvent évoluer dans l'espace comme dans le temps.
- Avoir un jeu de variables identiques entre prédictions de niches permet la comparaison de différentes populations vivant dans des régions différentes.

Selon Vaughan et Ormerod [2003], la capacité prédictive des algorithmes peut être améliorée en tenant compte des interactions biotiques et en sélectionnant les variables biologiquement reliées à l'espèce, spécialement lorsqu'il s'agit d'étudier des réponses environnementales aux changements climatiques sur de grandes étendues. Il est certain qu'un tel apport permettrait de mieux appréhender la niche réalisée de l'espèce (voir section 1.2.3), et faciliterait l'interprétation, mais les interactions biologiques sont difficiles à intégrer en tant que variables environnementales. Notre stratégie consiste à conserver un maximum de ces variables environnementales, au risque qu'il y ait une redondance. En effet, nous estimons

que la combinaison de ces dernières pourraient simuler, en arrière-plan, tout ou partie des interactions biotiques éventuelles. Ainsi, pour chaque modélisation effectuée dans cette thèse, nous conserverons le jeu de 26 variables environnementales (19 bioclimatiques tableau 2.6 et 7 topographiques tableau 2.5).

2.4.2.2 Réduire la zone de prédiction : établissement du M

La partie de l'environnement accessible par une espèce est dénommée M (voir section 1.2.4). Barve *et al.* [2011] ont démontré que la taille de cette zone pouvait avoir un impact sur la qualité de la prédiction d'une niche écologique. Définir le M est donc une tâche relativement délicate car le modélisateur doit restreindre l'étendue géographique de l'étude selon ses connaissances et selon ses *a priori* sur l'espèce, tout en prenant garde de ne pas masquer des zones ayant le potentiel écologique d'accueillir celle-ci. La mise en évidence de telles zones étant souvent l'enjeu des études, une tendance à surestimer l'aire du M, par exemple en assimilant son étendue à la totalité de la région étudiée, peut être observée. Ainsi, nombre d'études publiées ne définissant pas de M (par exemple : Holzmann *et al.* [2015]; Monroe *et al.* [2014]) sont conduites régulièrement. Lorsque nous cherchons à reconstruire les niches écoculturelles de populations du passé, nous sommes également confronté à cette situation, mais à moindre mesure, car la mobilité des groupes humains est accrue et évolue avec le temps grâce à des moyens de transports sans cesse innovants ainsi que des technologies permettant de surmonter des handicaps imposés par l'environnement. Ainsi pour modéliser la niche de populations de chasseurs cueilleurs, Banks *et al.* [2011] définissent un M s'étendant selon une zone tampon de 175 km autour des sites archéologiques utilisés pour l'étude, puis découpent chacune des couches environnementales avant d'effectuer l'ECNM. De même, Banks *et al.* [2013a] définissent un M couvrant une grande partie de l'Europe de l'Ouest ainsi que le Nord de l'Afrique avant de procéder à la modélisation des niches écoculturelles de populations dont les auteurs estiment qu'elles étaient dans la capacité technologique d'entretenir de vastes réseaux sociaux.

Les occurrences et les variables environnementales, soigneusement sélectionnées ou dont l'usage a été justifié, sont ensuite injectées dans des algorithmes prédictifs afin d'obtenir des modèles de niches, mais avant de poursuivre et d'analyser les résultats, il est possible d'optimiser les prédictions obtenues.

2.5 Optimiser une prédiction

Les modèles « bruts » : sorties originales des ENM, sont perfectibles. En effet on peut obtenir, par exemple, des prédictions qui comprennent une probabilité de présence non nulle sur l'ensemble d'une zone d'étude ou à l'inverse des valeurs de prédictions de présence très faibles pour des zones où l'espèce est réellement présente. Il est alors légitime

de se demander à partir de quelle valeur la prédiction de présence peut être considérée comme significative. Cette valeur peut varier selon la raison pour laquelle l'étude est menée : un seuil de valeur faible sera retenu pour éviter l'invasion d'espèces exotiques, pour la conservation ou pour la gestion du milieu, alors qu'à l'inverse une valeur forte sera choisie concernant la réintroduction d'espèces, la protection d'espèces en danger, ou dans notre cas, la détermination de zone appropriées à l'expression de traits culturels spécifiques.

Les modèles « bruts » se trouvent souvent peu informatifs car peu ou trop sensibles. Les prédictions d'un modèle peu sensible recouvrent une surface beaucoup plus importante que la réelle expression géographique de la niche de l'organisme étudié et, à l'inverse, les modèles très sensibles omettent de prédire des régions entières où l'espèce a la possibilité de se trouver. Lorsqu'on s'aperçoit qu'un modèle est trop sensible, on doit alors modifier les paramètres propres à l'algorithme utilisé afin d'obtenir des prédictions plus « larges ». Dans le cas d'une prédiction peu précise, car trop étendue, l'application d'un seuil sur les valeurs de la prédiction permet de réduire ce biais.

2.5.1 Utiliser plusieurs jeux de variables environnementales

Le choix des variables environnementales influence l'issue de la modélisation de niche [Peterson et Nakazawa, 2008]. Il est donc judicieux de mener différentes modélisations avec plusieurs jeux de variables environnementales pour une étude donnée. Cependant, nous souhaitons dans cette thèse conduire des modélisations dans plusieurs régions et à des époques différentes. Nous avons donc choisi d'homogénéiser la méthode en n'utilisant pour chacun des cas des variables bioclimatiques (tableau 2.6) issues d'une seule simulation climatique. Ce choix admet, certes un risque d'erreurs, mais il permet en revanche une comparaison et une meilleure interprétation des ECNM réalisées.

2.5.2 Les modélisations seuillées

Les écologues modélisateurs de niches ont proposé diverses méthodes permettant d'appliquer un seuil à une prédiction. Le choix de cette méthode s'effectue selon la nature des occurrences (présence/absence ou présence seulement) mais également en fonction de la qualité et de la fiabilité des occurrences. Selon la rigueur exigée par l'étude, la valeur de ce seuil est plus ou moins élevée ou astreignante et procure une valeur informative à la modélisation. L'attribution d'un seuil permet également d'homogénéiser les prédictions issues de différents algorithmes (ou des mêmes algorithmes, mais avec différents réglages ou différentes variables). Avoir des modèles homogènes offre alors la possibilité de comparer des niches écologiques de plusieurs espèces ou modélisées différemment. L'utilisation d'un seuil est également justifié afin d'éviter des biais induits par l'éventuelle présence de données erronées parmi les occurrences. L'emploi de seuil est donc très utile et la méthode d'application reste à l'appréciation du modélisateur.

2.5.3 Quels seuils appliquer en ECNM ?

Norris [2014] a démontré que la variabilité des prédictions pouvait être plus influencée par le choix du seuil que par la nature des occurrences. De nombreuses façons d'appliquer un seuil ont été décrites dans la littérature [Liu *et al.*, 2005; Peterson *et al.*, 2011], le tableau 2.5.3 résume différentes méthodes d'attribution de valeur à ce seuil, cette table ne liste que des méthodes qui s'appliquent aux modélisations produites à partir de données n'indiquant que la présence d'une espèce (*presence only*).

Tableau 2.7 : Différentes méthodes pour définir un seuil à appliquer à une prédiction. Ce seuil correspond à 1-E (où E représente l'erreur d'omission « tolérée »). Cette table présente une sélection de méthodes présentées par Hijmans *et al.* [2013]; Liu *et al.* [2005]; Peterson *et al.* [2011] (nous ne présentons ici que les méthodes concernant les modèles *presence only*).

Méthode	Définition	Référence
[†] <i>Fixed value</i>	Une valeur fixée arbitrairement (ex : 0.5)	[Manel <i>et al.</i> , 1999; Robertson <i>et al.</i> , 2004]
* [†] <i>Least Training Presence (LPT)</i>	La plus petite valeur de prédiction parmi les occurrences de présence	[Pearson <i>et al.</i> , 2007]
[†] <i>Fixed sensitivity</i>	Valeur correspondant à un pourcentage d'occurrences devant être prédites présentes	[Pearson et Dawson, 2004]
<i>Average probability/suitability</i>	La moyenne des probabilités de prédiction du modèle	[Cramer, 2003]
<i>Mid-point probability/suitability</i>	La médiane des probabilités de prédiction du modèle	[Fielding et Haworth, 1995]
* <i>Max Kappa</i>	Seuil pour lequel la valeur de kappa est la plus élevée	[Hijmans <i>et al.</i> , 2013]
* <i>Spec-sens</i>	Seuil pour lequel la somme de la sensibilité et de la spécificité est le plus élevé	[Hijmans <i>et al.</i> , 2013]
* <i>Equal Prevalence</i>	Seuil pour lequel la prévalence du modèle est la plus proche de la prévalence observée	[Cramer, 2003]
* <i>Equal-sens-spec</i>	Seuil pour lequel la sensibilité est égale à la spécificité	[Hijmans <i>et al.</i> , 2013]
<i>ROC plot-based approach</i>	Le seuil correspond à l'ordonnée du point de la courbe ROC le plus proche de l'angle du haut-gauche du graphique	[Cantor <i>et al.</i> , 1999]

(*) méthode de seuil pouvant être calculée avec le *package* R : *dismo* [Hijmans *et al.*, 2013].

(†) fonction redéveloppées sous R pour le besoin de la thèse voir

Parmi les seuils présentés tableau 2.5.3, deux sont fréquemment utilisés : *Least Training Presence* et *Fixed sensitivity*.

Least Training Presence. La valeur du seuil correspond à la valeur de la prédiction la plus basse parmi les occurrences de présence. Les valeurs en dessous de ce seuil sont réduites à 0.

Fixed sensitivity. Ce seuil est basé sur l'erreur d'omission (E) que le modélisateur définit au préalable (seuil fixé en général à 5% ou 10%). Après avoir classé les vrais positifs (voir figure 2.6) par ordre croissant des valeurs de prédictions de présence attribuées par le modèle, on retire les E% premiers, puis on établit comme valeur seuil la probabilité du premier non retiré. On affecte ensuite une valeur nulle en tout point de la prédiction ayant une valeur en dessous de ce seuil.

✍ ENM du lemming actuel : nous avons utilisé le seuil *Fixed sensitivity* avec $E = 0.05$. La valeur de prédiction la plus haute parmi les 5% les plus faibles des 204 occurrences correspond à la valeur attribuée à la 10^{ème} occurrence lorsque ces dernières sont classées par ordre croissant des prédictions que leur a attribué le modèle. Cette valeur étant de 0.4 on a donc fixé le seuil à cette hauteur, c'est-à-dire que nous avons affecté une probabilité de présence nulle à tous les pixels de la prédiction ayant une valeur inférieure à 0.4. Cette opération a été effectuée grâce à une fonction développée sous R.

Les écologues estiment qu'une prédiction seuillée est beaucoup plus informative qu'une prédiction brute, cependant cela ne la rend pas plus pertinente. Dans la section suivante, nous allons voir la variabilité qui peut exister entre différentes prédictions et comment les évaluer.

2.6 Stochasticité, validité et pertinence des prédictions

Chaque algorithme ayant ses spécificités, les prédictions qu'ils fournissent sont très variées. Ces différences sont aussi observables parmi les prédictions fournies par un même algorithme, car fonctionnant par apprentissage ils laissent place à une part d'aléatoire à chaque exécution. Cette approche stochastique permet aux algorithmes d'explorer l'ensemble des solutions puis, par des méthodes statistiques, de ne conserver que les plus pertinentes. Afin d'optimiser l'exploration de solutions, il est également possible de paramétrer les algorithmes de façon très différentes afin d'avoir des algorithmes très sensibles ou au contraire permettant une grande flexibilité (nous qualifions d'« optimiste », ce deuxième paramétrage).

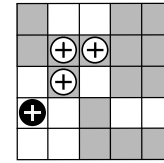
Face à la diversité intra ou inter-algorithmes des prédictions ont été développées des méthodes basées sur la matrice de confusion permettant d'évaluer leur pertinence.

2.6.1 La matrice de confusion

Cet outil d'évaluation de modèle présente en colonnes les effectifs des classes d'occurrences (présences et absences réelles de l'espèce) et en lignes le nombre de prédictions d'absence et de présence. Ces tests permettent d'évaluer la fiabilité des modèles ainsi que la comparaison entre des modèles obtenus suivant différentes méthodes (issues d'algorithmes différents ou semblables mais avec des paramétrages distincts).

La figure 2.6 montre l'ensemble des états possibles que peut prendre un point (c'est-à-dire un lieu donné, un pixel) d'une carte reflétant la projection géographique d'un modèle de niche. La prédiction peut s'avérer juste dans deux cas : pour un lieu où l'espèce est prédite présente et est réellement présente (vrai positif) ainsi que pour un lieu où l'espèce, prédite absente est réellement absente. De la même façon, les modèles comportent deux types d'erreurs : l'erreur d'omission où un point d'occurrence n'a pas été prédit comme un lieu de présence et l'erreur de commission où une présence est prédite en un endroit où

Prédictions (modèles)	Observations (occurrences)	
	présence	absence
	présence	absence
	a	b
	c	d



- (a) **Vrais positifs** : présences du jeu de données initial pour lesquelles le modèle prédit une présence
- (b) **Faux positifs** : points prédits présents ne faisant pas partie des occurrences
- (c) **Faux négatifs** : présences initiales pour lesquelles le modèle prédit une absence
- (d) **Vrais négatifs** : prédictions d'absence qui ne sont pas des occurrences

Les carrés gris foncés correspondent à la niche écologique prédite - les cercles représentent les occurrences, blanc : vrai positif (a), noir : faux négatif (b) - les cases grises (excepté celles où il y a des occurrences) sont des faux positifs (b) - les cases blanches (à l'exception de celles contenant une présence) représentent les vrais négatifs.

FIGURE 2.6 : Matrice de confusion et représentation spatiale schématique des différents états d'un point dans un modèle.

l'espèce ne se trouve pas. Dans la matrice de confusion, les cellules a et d correspondent aux prédictions exactes du modèle, la case b à l'erreur de commission et la case c à l'erreur d'omission.

2.6.2 Évaluer une prédiction

Confronté à un modèle de niche, il est important pour le modélisateur de connaître sa pertinence, c'est-à-dire d'apprécier l'exactitude de la prédiction. Pour cela il existe plusieurs tests basés sur la matrice de confusion (figure 2.6). La grille à la droite de la figure est la schématisation d'une carte où apparaît la niche écologique prédite et les occurrences d'espèce, elle montre comment se manifestent dans l'espace géographique les quatre situations précédemment décrites. On voit ainsi que l'erreur de commission correspond en réalité aux prédictions de présence recherchées, par conséquent cette erreur doit être contrôlée afin d'obtenir un modèle le plus correct possible. En effet, si cette erreur est trop importante, alors la niche écologique prédite est trop grande et à l'inverse, si le taux de commission est trop faible, alors la niche se trouve restreinte à une zone avoisinant les points d'occurrences et la prédiction dans ce cas est peu informative. Toutes les prédictions de présence d'un modèle représentent la niche écologique de l'espèce étudiée, soit la somme des vrais positifs et des vrais négatifs (équation 2.1).

$$\text{Niche}_{\text{prédite}} = a + b \quad (2.1)$$

À partir des occurrences qui attestent de la présence réelle de l'espèce et des prédictions des algorithmes, on peut estimer la validité intrinsèque des modèles en mesurant leur sensibilité et leur spécificité.

2.6.2.1 La sensibilité

La sensibilité (équation 2.2) est la probabilité que la présence soit prédite en un lieu où elle est attestée, elle se mesure uniquement parmi les présences. Cette mesure équivaut à la proportion de sites, réellement présents et prédits présents par le modèle.

$$\text{sensibilité} = \frac{a}{a + c} \quad (2.2)$$

Un modèle « parfait » est un modèle qui ne fait pas, entre autres, d'erreur d'omission. Un tel modèle possède alors une sensibilité optimale. Ceci caractérise un modèle pertinent qui prédit une présence en tous lieux convenables à l'espèce donc dans sa niche écologique. Typiquement un modélisateur va chercher à minimiser les omissions en forçant les algorithmes à ne fournir que des modèles dépourvus d'erreur d'omission. Cependant considérant les impacts qu'ont sur les prédictions de niches les erreurs d'échantillonnage [Peterson *et al.*, 2014] (issues, par exemple d'un mauvais relevé GPS ou d'une mauvaise détermination) ainsi que la provenance des occurrences (zone « puits », ou zone « source », voir la dynamique source-puits tableau 1.1), il est préférable de ne pas fixer cette valeur à 0, afin que l'algorithme puisse classer des relevés de présence comme en dehors de leur niche écologique.

2.6.2.2 La spécificité

La sensibilité ne doit pas être prise en compte sans regarder la spécificité (équation 2.3). En effet un modèle qui prédirait uniquement les occurrences de présence aurait une sensibilité de 1 or nous cherchons en ENM à prédire d'autres pixels que ceux où la présence est attestée par les données d'occurrences. La spécificité mesure la proportion d'absences correctement prédites, soit la probabilité que le modèle prédise une absence où l'espèce à préalablement été présumée absente.

$$\text{spécificité} = \frac{d}{b + d} \quad (2.3)$$

La spécificité quantifie l'erreur de commission $\text{Err}_{\text{commission}} = \frac{b}{b+d}$. En ramenant $b + d$ à 1 (100% des absences) on a alors $\frac{d}{b+d} + \frac{b}{b+d} = 1$ d'où $\text{Err}_{\text{commission}} = 1 - \text{spécificité}$

En ECNM nous n'utilisons pas les données d'absence dans ce cas, pour procéder à l'évaluation par la spécificité, des pseudos absences sont utilisées.

2.6.2.3 Les pseudos absences

Pour pallier le fait que les relevés sont souvent des relevés de présence uniquement, particulièrement en archéologie (section 2.6.2.4), on puise les absences aléatoirement parmi des pixels de l'environnement pour lesquels une présence n'a pas été signalée. Ainsi b devient estimable et on peut en déduire la spécificité du modèle qui, alliée à la sensibilité,

indiquent sa performance. Cela implique une sous estimation de l'erreur de commission car il est possible que l'absence provenant de l'échantillonnage aléatoire soit un faux négatif.

2.6.2.4 Le taux d'omission en archéologie

Lors des prospections d'espèces actuelles, il advient qu'une espèce soit indiquée présente dans un habitat qui ne soit pas le sien, l'individu pouvant se retrouver isolé pour différentes raisons. Dans ce cas, la répartition de l'échantillon ne reflète pas l'écologie de l'espèce. On admet alors un taux d'omission dans les modèles (plus ou moins intense en fonction de la qualité de l'échantillonnage, du bruit engendré par l'anthropisation des milieux ou autres perturbations). En archéologie la présence de plusieurs vestiges est nécessaire pour prouver l'appartenance d'un site à une entité culturelle. De plus, la probabilité de retrouver des vestiges, dont la présence hors de leur aire culturelle serait fortuite, est fortement abaissée par le fait que seulement une infime partie de cette aire puisse être fouillée. En ce sens, la présence d'objets caractéristiques étant un fait archéologique, on doit considérer comme fausses les prédictions classant des sites hors de leur aire culturelle mais contenant pourtant de tels vestiges. Nous ne retiendrons donc que les prédictions de niches ayant un taux d'omission nul.

2.6.3 Tests de performance des modèles

Nous venons de voir dans la section précédente que la sensibilité et la spécificité nous informaient quant à la performance d'un modèle, seulement il est difficile d'apprécier si un modèle plus sensible ou plus spécifique doit être privilégié et les importances que l'on accorde respectivement aux erreurs de commission et d'omission. Dans la section suivante nous décrivons des tests permettant d'évaluer les modèles en combinant ces caractéristiques.

2.6.3.1 Accuracy et prévalence

L'*accuracy* (ou la précision) est la proportion d'occurrences correctement classées par le modèle (équation 2.4), alors que la prévalence (équation 2.5) est la proportion de présence sur l'ensemble des sites [Allouche *et al.*, 2006]. Lors d'une modélisation cette deuxième proportion est généralement faible car le nombre d'occurrences, aussi important soit-il, reste toujours très faible par rapport à la quantité de sites prédits c'est-à-dire au nombre de pixels de la zone étudiée (par exemple 100 occurrences sur une zone de 10 000 km² avec un maillage de 1 km² donne une prévalence de 0.01)

$$\text{accuracy} = \frac{a + d}{n} \quad \text{où } n = a + b + c + d \quad (2.4)$$

$$\text{prévalence} = \frac{a + c}{n} \quad \text{où } n = a + b + c + d \quad (2.5)$$

2.6.3.2 Cas particulier de la prédiction en présence seulement

Lorsqu'on effectue une modélisation uniquement avec des occurrences de présence, on ne dispose pas de la matrice de confusion entière. En effet, il est impossible de discerner les faux positifs (b) des vrais négatifs (d), or les faux positifs représentent la prédiction. C'est précisément la pertinence de ces derniers que nous cherchons à évaluer. La spécificité ne peut pas être évaluée mais on peut cependant l'estimer en fonction de la proportion de faux positifs.

Plusieurs mesures pour estimer la performance des modèles dérivent du rapport sensibilité/spécificité, de la précision ou de la prévalence.

2.6.3.3 κ de Cohen

Le test κ de Cohen dit Kappa [Cohen, 1960] est la statistique la plus utilisée pour mesurer la performance des modèles générés à partir d'occurrences de présence-absence [Allouche et al., 2006]. Cette statistique cherche à estimer l'*accuracy* obtenue (P_A) par le modèle en la comparant à l'*accuracy* qu'on obtiendrait si les prédictions avaient été produites de manière aléatoire (P_E), voir équation 2.6.

$$\kappa = \frac{P_A - P_E}{1 - P_E} \quad (2.6)$$

La valeur de κ fluctue entre -1 et 1, une valeur proche de 1 indique un modèle pertinent alors qu'une valeur nulle indique un tirage aléatoire, et une valeur proche de -1 une corrélation inversée. Allouche et al. [2006] a exprimé le κ en fonction de la matrice de confusion équation 2.7

$$\kappa = \frac{\left(\frac{a+d}{n}\right) - \frac{(a+b)(a+c)+(c+d)(d+b)}{n^2}}{1 - \frac{(a+b)(a+c)+(c+d)(d+b)}{n^2}} \quad (2.7)$$

En dépit de sa large utilisation, plusieurs études ont montré que la statistique kappa était biaisée intrinsèquement du fait de sa forte dépendance à la prévalence, ce qui la rend inefficace pour estimer la performance d'un modèle [Allouche et al., 2006].

2.6.3.4 TSS

Le TSS (en anglais *true skill statistic*, équation 2.8) conserve les avantages du kappa tout en étant indépendant de la prévalence (pour les modèles de présence-absence). Ce test tient compte des erreurs d'omission et de commission ainsi que de la différence avec un modèle aléatoire. La valeur du TSS est comprise entre -1 et 1, où 1 indique une égalité parfaite et 0 ou moins une performance moins bonne que celle d'un modèle aléatoire.

$$\text{TSS} = \text{sensitivité} + \text{spécificité} - 1 \quad (2.8)$$

2.6.3.5 AUC

À l'instar des études cliniques dans le domaine de la santé destinées à mesurer la pertinence de diagnostics de maladies, de nombreuses études d'ENM utilisent des courbes de sensibilité/spécificité appelées courbes ROC (*Receiver Operating Characteristics* [DeLong *et al.*, 1988], figure 2.7) afin d'estimer le potentiel prédictif des algorithmes. La courbe ROC, indépendante de la prévalence [Allouche *et al.*, 2006], estime la spécificité (absence d'erreur de commission) et la sensibilité (absence d'erreur d'omission) en traçant le taux de vrais positifs en fonction du taux de faux positifs d'un modèle. Cette courbe permet ainsi d'évaluer la performance de celui-ci. De cette manière, on estime la pertinence du modèle établi en mesurant l'aire se trouvant sous cette courbe. Plus la partie convexe de la courbe se trouve vers le coin supérieur gauche du graphique, autrement dit plus l'aire sous la courbe ROC est grande, plus le potentiel prédictif du modèle est fort [Fielding et Bell, 1997; Zweig et Campbell, 1993]. Cette aire, à la surface comprise entre 0 et 1, est nommée AUC (pour *Area Under the Curve*). On considère une prédiction comme due au hasard si l'AUC équivaut à 0.5 et comme forte dès lors que l'AUC dépasse 0.75 (remarque : si un modèle a une AUC inférieure à 0.25 il suffira de prendre l'inverse de sa prédiction pour avoir un modèle avec une forte probabilité d'être juste). Cette approche confère l'avantage d'être indépendante à la prévalence [Allouche *et al.*, 2006].

L'AUC est souvent utilisée comme seule mesure pour évaluer la performance des modèles [Allouche *et al.*, 2006] seulement le rapport vrais positifs/faux positifs est à considérer avec réserves, car nous verrons dans le chapitre 6 que deux prédictions de niches distinctes peuvent avoir la même valeur d'AUC. Plusieurs études ont signalé des problèmes avec cette approche [Lobo *et al.*, 2008; Peterson *et al.*, 2008a; Raes et ter Steege, 2007], cités par [Araújo *et al.*, 2014].

Basé sur la courbe ROC également une autre mesure d'aire : le *partial ROC ratio*, permet une meilleure identification de la pertinence des modèles.

2.6.3.6 Partial ROC ratio

Peterson *et al.* [2008a] introduisent le calcul du *partial ROC ratio* [Dodd et Pepe, 2003; Jiang *et al.*, 1996] en ENM. Cette statistique mesure l'aire sous la courbe ROC dans une région particulière qui est l'intersection de deux intervalles figure 2.8. Chacun de ces intervalles sont reportés sur un axe du graphique :

- Le premier intervalle se situe sur l'axe des abscisses et représente donc une portion des faux positifs. Il correspond à l'ensemble des valeurs de prédictions attribuées par l'algorithme prédictif. Cet intervalle est borné entre la valeur de prédictions minimale et maximale.
- Le second, sur l'axe des ordonnées, est relatif aux vrais positifs, il s'agit de l'intervalle de tolérance à l'erreur d'omission (c'est-à-dire l'erreur d'omission que l'on est prêt

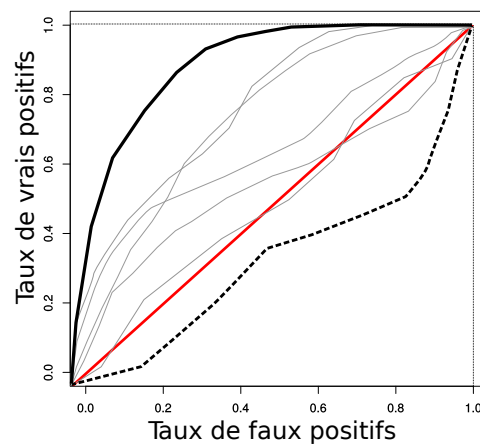


FIGURE 2.7 : L'AUC, aire en dessous de la courbe ROC. L'AUC permet d'évaluer la pertinence d'un modèle. La courbe ROC est la courbe des vrais positifs en fonction des faux négatifs, elle peut être aussi exprimée comme étant la courbe de sensibilité (1 - taux d'omission) en fonction de 1-spécificité (la proportion d'occurrences de présence prédites). Plus l'AUC est grande (c'est à dire proche de 1) plus le potentiel prédictif du modèle est grand. Une AUC de 0.5, représenté par l'aire en dessous de la droite passant par l'origine et le point de coordonnées (1,1), correspond à une classification totalement aléatoire, en dessus de cette droite la probabilité que les prédictions du classificateur soient justes augmente et à l'inverse diminue en dessous de cette dernière. La courbe noire en gras représente une bonne prédiction, la courbe en pointillés une mauvaise prédiction et la droite rouge une prédiction aléatoire. Remarquons que dans le cas d'une AUC très faible on obtiendra une bonne classification en inversant la prédiction.

à accepter, cette erreur est égale à E). Ce dernier est borné par le point de la courbe correspondant à $1-E$ et 1 (la sensibilité maximale).

L'intersection de ses deux intervalles peut être reportée sur l'axe des abscisses (représentant 1-spécificité) et bornée entre deux valeurs que nous nommons α et β . Afin d'estimer la part de la prédiction qui n'est pas aléatoire nous faisons le rapport de l'aire sous la courbe ROC et sous la droite du modèle nul dans l'intervalle $[\alpha; \beta]$.

Lorsque le *partial ROC ratio* a une valeur avoisinant 1, alors la prédiction est proche du modèle nul. Une valeur inférieure à 1 signale un modèle qui a plus de chance de se tromper que ne le ferait un modèle nul. On estime qu'une valeur de *partial ROC ratio* supérieure à 1.5 indique une bonne prédiction.

Les statistiques *accuracy*, Kappa et TSS nécessitent des données de types absence. Ainsi, seules les statistiques issues de la courbe ROC peuvent servir à évaluer les prédictions avec pertinence. Toutefois, [Peterson et al. \[2008a\]](#) signalent que l'évaluation des prédictions par l'AUC confère des avantages aux algorithmes macroprédictifs. Il est à noter également que le *partial ROC ratio* est plus en congruence avec l'écologie que l'AUC car il permet l'évaluation

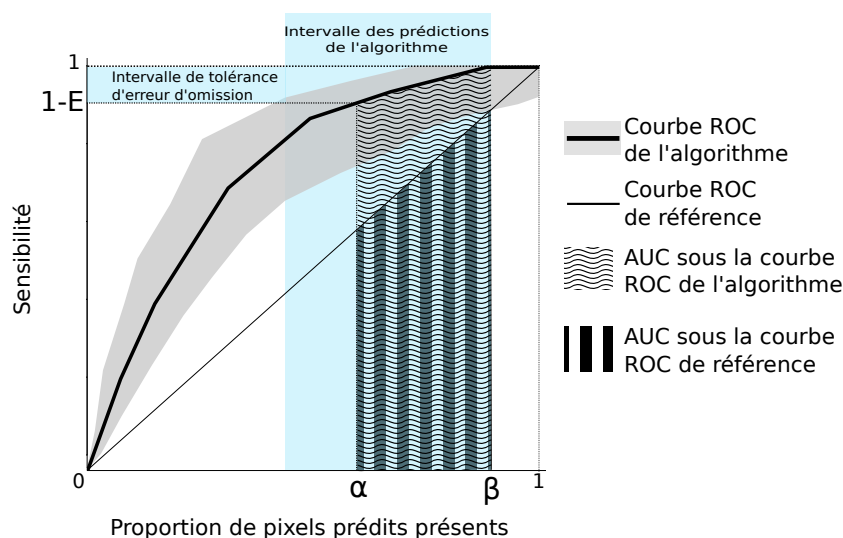


FIGURE 2.8 : Représentation graphique du partial ROC ratio qui permet d'évaluer la pertinence d'un algorithme. L'aire grisée représente l'ensemble des courbes ROC pour chacune des prédictions de l'algorithme évalué, la courbe noire au centre de cette aire représente la courbe ROC moyenne de ces prédictions. Le partial ROC ratio est la division entre les aires situées sous la courbe ROC de l'algorithme à évaluer et sous la droite de référence $y = x$ (qui correspond à un modèle nul, c'est-à-dire fournissant des prédictions aléatoires). Ces aires sont réduites à l'intersection des intervalles de tolérance d'erreur d'omission et celui des valeurs de prédictions de l'algorithme $[\alpha; \beta]$.

de la pertinence de l'algorithme dans la région spécifique de la courbe ROC correspondant à l'intersection de l'intervalle de prédiction et l'intervalle d'erreur d'omission (figure 2.8).

Basé sur les proportions de faux négatifs et de faux positifs des prédictions issues de plusieurs algorithmes peuvent, bien que distinctes, avoir été évaluées comme pertinentes. N'ayant pas la capacité de juger quel est la meilleure solution, il peut donc s'avérer judicieux de considérer l'ensemble des résultats. Nous verrons (chapitre 6) comment nous nous servons des statistiques présentées ici pour sélectionner et pour pondérer les prédictions avant de les regrouper dans un consensus.

Synthèse du chapitre : Les algorithmes prédictifs de niches écologiques

Ce chapitre récapitule l'utilisation et le fonctionnement des algorithmes prédictifs de niches.

Après une brève énumération des possibilités qu'offre la modélisation de niche en écologie et en archéologie, nous avons défini les spécificités que requière l'ECNM, et conclu qu'il est préférable d'établir des prédictions de niches avec des algorithmes pouvant se passer de données d'absence.

Il a été abordé également le caractère stochastique des prédictions dû à l'hétérogénéité des algorithmes prédictifs et de ce fait nous avons souligné la nécessité d'évaluer ces prédictions.

Après avoir présenté plusieurs tests statistiques nous concluons que seules les statistiques basées sur la courbe ROC peuvent être appliquées pour juger de la pertinence d'une prédiction obtenue à partir d'occurrences de présence uniquement. Parmi ces dernières, le *partial ROC ratio* apparaît comme le plus informatif.

Enfin, nous avons vu la nécessité d'appliquer un seuil aux prédictions.

DEUXIÈME PARTIE

DÉVELOPPEMENT D'OUTILS SPÉCIFIQUES À LA MODÉLISATION DE NICHES

Collecter les occurrences

Préambule

Ce chapitre décrit les méthodes employées dans cette thèse pour prélever des occurrences de présence de façon automatique ou semi-automatique grâce à des requêtes spatiales.

Sommaire

3.1	Méthodes de collectes classiques	85
3.2	Affiliation à partir de régions prédéfinies	86
3.3	Extractions d'occurrences à partir de cartes scannées	87

3.1 Méthodes de collectes classiques

Afin de procéder à une ENM ou à une ECNM, les modélisateurs peuvent recueillir les données d'occurrences en prélevant eux-mêmes les données sur le terrain (données d'observations), en exploitant les bases de données (muséales, publiées, ou disponibles en ligne) ou en les récoltant dans la littérature scientifique. Ces collectes nécessitent de longues périodes d'investigations et beaucoup de rigueur. Bien qu'un échantillon d'occurrences de l'ordre de la vingtaine d'individus soit suffisant pour une modélisation de niche (voir section 2.3.1.1), ce dernier doit être suffisamment représentatif de l'organisme étudié. Pour atteindre des données suffisamment diversifiées, la taille de l'échantillon nécessaire peut donc considérablement dépasser la taille minimale requise par les algorithmes prédictifs. Atteindre un tel effectif peut s'avérer périlleux lorsqu'il s'agit par exemple de données difficiles à collecter sur le terrain ou encore de données très rares comme les celles issues du registre fossile.

Les modélisations de niches que nous menons sur des groupes ethnolinguistiques (chapitre 8) ou sur des sociétés historiques (chapitre 9) sont à vocations référentielles. Les occurrences que nous sélectionnons pour ces modélisations doivent par conséquent représenter le maximum de la variabilité des paramètres environnementaux des niches écoculturelles. Dans cette perspective, nous devons multiplier considérablement les données d'occurrences. Ainsi, nous avons mis au point deux techniques de collecte de données d'occurrences applicables lorsque les listes de coordonnées géographiques associées aux catégories étudiées n'existent pas ou n'ont pas été publiées.

3.2 Affiliation à partir de régions prédéfinies

La technique que nous décrivons permet de classer une liste de localités préétablies de façon automatique selon des cartes recueillant des informations culturelles. Pour mener cette opération, nous devons posséder la liste des coordonnées géographiques des points dont on souhaite affecter une attribution automatiquement et des cartes de ces attributions. Le format des cartes importe peu, l'unique condition requise est qu'elles aient au moins trois repères spatiaux relativement bien identifiables comme des éléments géographiques distinguables (pic, presque-île, ville, monument, etc) ou un quadrillage, une échelle et une orientation. Ici, nous détaillons la manière dont nous avons procédé, pour collecter les données d'occurrences qui ont servi à reconstruire les niches ecolinguistiques de Nouvelle-Guinée (chapitre 8). Il s'agissait de recouper l'information apportée par une carte des groupes ethnolinguistiques d'une part et d'une liste de villages d'autre part (voir section 8.5.1).

La procédure est la suivante :

1. Création de fichiers images séparés de mêmes dimensions et mêmes résolutions correspondant chacun à une aire linguistique (nous appelons aire linguistique, l'aire attribuée à un groupe de langue). Cette opération est réalisée à l'aide d'un *script-fu*, langage de programmation pour générer des macros sous le logiciel *The Gimp* (<http://gimp.org>). Ce script est capable de séparer les éléments de couleurs différentes d'une image.
2. Géoréférencement de l'une des images générées à l'étape précédente grâce au logiciel *qGIS* [Team, 2009]
3. Affectation du géoréférencement de chacune de ces images grâce une boucle de programmation réalisée sous R
4. Conversion automatisée en ascii (fichier texte) de chacune des images grâce au *package Raster* de R
5. Utilisation d'une boucle de programmation sur chacun des fichiers ascii géoréférencés. La boucle permet de retrouver l'affectation de chaque point de la liste en fonction de sa position géographique.

À l'issue de ce procédé, nous obtenons une affectation pour chaque point se trouvant dans les polygones de la carte de référence. Cette liste peut alors être utilisée pour l'ECNM.

Cette manipulation est très utile lorsque les sources de données proviennent de cartes puisées dans la littérature présentant des échelles ou des systèmes de projections différents ou lorsque la liste des points à identifier est longue. La procédure procure des avantages, mais aussi des biais dont il faut avoir connaissance (tableau 3.1).

Tableau 3.1 : Avantages et désavantages de l'affiliation et de la collecte d'occurrences automatique

Points forts	Points faibles
Permet de recueillir beaucoup d'informations pour les occurrences et constituer des groupes en croisant ces informations.	La précision dépend fortement de la qualité des cartes de références.
Permet de recueillir des informations à partir de nombreuses sources malgré des échelles différentes.	Un travail laborieux de traitement d'image au préalable peut être requis.
Évite les erreurs dues à un travail manuel (inversion de coordonnées, erreurs de saisies).	L'affectation des occurrences en bordure des polygones est dépendante de la résolution des cartes sources.
Rapide et permet l'obtention de nombreuses occurrences par catégories.	
La réactualisation ou la modification de la requête demande peu de temps.	

3.3 Extractions d'occurrences de présence à partir de cartes scannées

Nombre d'auteurs publient des relevés de présence d'espèces ou de sites archéologiques sous forme cartographique sans jamais fournir les coordonnées géographiques. Il est possible alors de mener sa propre recherche afin de connaître la localité précise du site et ainsi obtenir ces coordonnées. L'on peut aussi rechercher leurs positions exactes sur des photos aériennes ou, si l'on ne nécessite pas d'une grande précision, relever les coordonnées de la commune la plus proche. Pour les besoins de cette thèse, nous avons collecté les coordonnées géographiques de près de 500 occurrences issues de cartes très précises (figure 3.1). L'opération a été effectuée sous *qGIS*, ce logiciel nous a permis de géoréférencer chacune des cartes collectées et de placer une entité (un point) au-dessus

Obtenir des couches environnementales haute définition pour une période spécifique

Préambule

En ENM et en ECNM les modélisateurs utilisent des simulations climatiques à haute résolution. Or, la disponibilité de celles-ci limite leurs études à certaines régions ou à des époques prédéterminées. Ce chapitre présente une méthode statistique permettant d'obtenir des simulations à haute résolution focalisées sur des régions et ciblant des époques particulières à partir de simulation globale à basse résolution. Cette méthode peut être également appliquée pour obtenir des simulations climatiques correspondant à plusieurs périodes successives et ainsi permettre d'apprécier l'évolution spatio-temporelle d'une niche.

Sommaire

4.1	Modèles climatiques globaux	89
4.2	Modèles climatiques régionaux	90
4.3	Downscaling	90
4.4	Descente d'échelle statistique par modèle additif généralisé	92
4.5	Critiques de la méthode présentée	104

4.1 Modèles climatiques globaux

Un modèle climatique global est composé de couches de données basse résolution (maillage de 2-3° en latitude et longitude pour les plus précis¹) couvrant le monde entier.

1. De l'ordre de la centaine de kilomètres.

Une couche correspond aux valeurs d'un paramètre modélisé en chaque point de grille grâce à différents forçages (section 2.3.3.1). Le nombre de couches est fonction du nombre de variables, du pas de temps² et de la plage temporelle de la modélisation. Il est donc envisageable d'obtenir *a posteriori* une couche de données pour toute période souhaitée sur la plage de temps que couvre un modèle climatique. Seulement la couche obtenue sera à la résolution du modèle.

En ECNM ou en ENM l'obtention de données environnementales à très haute résolution est un enjeu important, car comme nous l'avons vu précédemment, la résolution des couches environnementales détermine la résolution des prédictions de niches. De plus, la nature des prédictions elle-même peut être affectée par la résolution des variables environnementales car les mécanismes écologiques se déroulant à petite échelle peuvent ne pas être perçus [Guisan et Thuiller, 2005]. C'est pourquoi les modélisateurs s'efforcent à se procurer des couches de données environnementales à très haute résolution. Il est possible d'obtenir des couches de plus fine résolution, soit à partir de modèles régionaux, soit grâce à des techniques d'interpolation.

4.2 Modèles climatiques régionaux

Pour réaliser une modélisation de niche paléoécologique ou écoculturelle d'une population du passé, il faut se procurer des couches environnementales correspondant à la région et à l'époque de l'espèce ou de la culture concernée. Ces couches environnementales peuvent provenir d'un zoom d'un modèle climatique pour un lieu et une époque donnée. Par exemple, Banks *et al.* [2013b] effectuent des ECNM avec des couches environnementales pour deux périodes du Paléolithique (~ 40 ka) extraites du modèle LMDZ3.3 de l'IPSL [Jost *et al.*, 2005] pour l'Europe à une résolution de 50 km. De telles extractions doivent être programmées à l'avance. Les simulations zoomées restent relativement rares et ne fournissent pas toujours des données qui correspondent aux problématiques et axes de recherches des écologues, paléontologues ou archéologues. Toutefois, certaines périodes suscitent plus d'intérêts et se trouvent par conséquent plus disponibles. Il s'agit du Dernier Interglaciaire (~120 ka BP), du Dernier Maximum Glaciaire (21 ka BP) ou de l'Holocène moyen (6000 BP).

Il est toutefois possible d'obtenir des couches climatiques à haute résolution en faisant une interpolation d'un modèle global.

4.3 Downscaling

Le *downscaling* est un terme anglais pouvant être traduit par descente d'échelle. Il permet d'obtenir depuis une simulation basse résolution (LD), une simulation à haute

2. Intervalles de temps entre chaque calcul, ex : les pas de temps pour les modèles de l'IPSL sont de 30 minutes.

Tableau 4.1 : Liste des données environnementales fournies par *WorldClim*. Le *downscaling* pour les données actuelles est construit à partir d'observations les autres ont pour origine des simulations climatiques globales.

Période	Résolution(s)
2070 (2061-2080 AD)	30", 2.5', 5', 10'
2050 (2041-2060 AD)	30", 2.5', 5', 10'
Actuel (1950-2000 AD)	30", 2.5', 5', 10'
Holocène moyen (6000 BP)	30", 2.5', 5', 10'
Dernier maximum glaciaire (22000 BP)	2.5', 5', 10'
Dernier interglaciaire (120000 - 140000 BP)	30"

résolution (HD) pour des régions et des périodes spécifiques. Les techniques de *downscaling*, appliquent le principe d'actualisme, partant de l'hypothèse que les mécanismes climatiques actuels pour une région donnée restent sensiblement les mêmes selon les époques. Ainsi, des données actuelles sont utilisées pour calibrer la modélisation (c'est-à-dire établir les relations variables expliquées et variables explicatives³).

Comme nous l'avons vu section 2.3.3.2, il existe plusieurs types de *downscaling*, mais ces derniers sont très peu utilisés par les modélisateurs de niches dans l'objectif de produire leur propre jeu de variables environnementales, car leur emploi nécessite des connaissances en climatologie afin de choisir les bons prédicteurs ainsi que des manipulations de fichiers à l'aide de logiciels spécialisés. De ce fait, les couches de données environnementales HD directement disponibles sont les plus utilisées, c'est le cas des données *WorldClim* que l'on peut télécharger à l'adresse suivante <http://worldclim.org/>. Sur ce site *Hijmans et al.* proposent des jeux de données bioclimatiques aux résolutions 10', 5', 2.5' et 30" (voir tableau 2.6 pour une description de ces variables). Ces couches ont été produites par un *downscaling* statistiques « calibré » avec des données du climat actuel et « appliqué » à différents modèles (voir site web *WorldClim* pour une liste complète). Ces variables bioclimatiques conviennent très bien à la modélisation de niches et sont d'ailleurs largement utilisées comme telles (exemples : [Bernardo-Silva et al., 2012; Lim et al., 2011; Peterson et al., 2007]). Or, ces simulations ne sont disponibles que pour une poignée de périodes, en effet le projet de recherche *WorldClim* fournit des variables correspondant seulement aux climats de six périodes (une actuelle, deux fin du XXI^e siècle et trois du passé, voir tableau 4.1).

Toutefois, il est possible de se procurer des couches de données provenant de *downscaling*, mais elles ciblent souvent des lieux et des époques spécifiques. De plus ces variables à la résolution augmentée ne sont pas nécessairement bioclimatiques ou ne permettent pas de les recalculer. En effet les variables bioclimatiques peuvent être déduites des données mensuelles des températures minimales, des températures maximales et des précipitations. Or, il est difficile de comprendre l'évolution des températures extrêmes dans un contexte de changement climatique car les facteurs dont elles dépendent sont différents de ceux dont

3. Correspondant aux termes anglais *predictand* et *predictors*.

dépendent les températures moyennes pour lesquelles le *downscaling* est éprouvé [Vrac *et al.*, 2007]. Dans ce cas, cette technique reste encore à perfectionner.

Conscients de ces limites et compte tenu de nos besoins, nous avons décidé de reconstruire, pour cette thèse, des variables bioclimatiques de deux périodes du dernier millénaire en utilisant une démarche statistique détaillée dans la section suivante.

4.4 Descente d'échelle statistique par modèle additif généralisé

Dans cette section nous décrivons la méthode statistique employée pour obtenir les couches de données bioclimatiques HD des régions de l'Atlantique Nord à partir d'une simulation [Servonnat, 2010] couvrant le dernier millénaire pour deux périodes qui correspondent à l'optimum médiéval climatique (OCM pour la période allant de 1000 à 1300 AD) et au début du petit âge glaciaire (PAG pour la période allant de 1301 à 1450 AD). À cet effet, nous avons appliqué une régression GAM⁴ [Hastie et Tibshirani, 1990] sur les températures minimales (tn), maximales (tx) et sur les précipitations (pr). Ces variables HD, ainsi obtenues, ont servi par la suite à la reconstruction des niches écoculturelles des Vikings des régions subarctiques (voir chapitre 9).

4.4.1 La modélisation GAM

Pour augmenter la résolution des trois variables citées précédemment, nous nous sommes inspirés du *downscaling* statistique développé par Levvasseur [2012] qui utilise des modèles additifs généralisés : GAM. La régression GAM permet de capturer les formes des relations entre une variable expliquée et plusieurs variables explicatives sans que l'on ait eu besoin de déterminer de formes *a priori*, donc sans influencer la sortie du modèle. De plus, GAM élargi les possibilités des GLM (régressions linéaires multiples) en permettant aussi bien des lissages paramétriques que non paramétriques. La modélisation GAM s'effectue en deux temps, une première étape constituant un « apprentissage » et une seconde étape qui est la « prédiction ». La première phase établit les relations entre les variables explicatives et un jeu de données puis la seconde applique le modèle réalisé par l'apprentissage à un autre jeu de données. En *downscaling*, on parle de « prédictors » en ce qui concerne les variables dépendantes ou explicatives et de « prédictants » en ce qui concerne les variables indépendantes ou expliquées. Nous utiliserons ces termes pour la suite de ce document.

La régression GAM va être effectuée sur deux simulations : la simulation de référence pour l'apprentissage et la simulation dont on veut augmenter la résolution.

4. *Generalized Additive Model*

Tableau 4.2 : Variables de la simulation CRU (*Climate Research Unit*).

Nom	Description	Unitée
x	Longitude	°(décimaux)
y	Latitude	°(décimaux)
pr	Précipitations	mm.mois ⁻¹
pr_cv	Précipitations convectives	%
rd0	Nombre de jours humides (>0.1mm)	jours
tmp	Température de l'air	°C
dtr	Variation des températures du jour	°C
reh	Humidité relative	%
sunp	Ensoleillement	%
frs	Nombre de jours où le sol est gelé	jours
elv	Altitude	m
wnd	Vent	m.s ⁻¹
cld	Nébulosité	%

4.4.1.1 La simulation de référence

La simulation CRU (*Climate Research Unit*), est une simulation HD [New *et al.*, 2002] du climat terrestre de tous les continents hormis l'Antarctique. Elle inclut les variables énumérées dans le tableau 4.2. Ces variables proviennent d'une interpolation d'enregistrements de stations entre janvier 1961 et décembre 1990. CRU présente une grille de 2160×1080 qui lui confère une résolution régulière (en longitude comme en latitude) de $1/6^\circ$ (soit $10'$). Cette simulation à très haute résolution (CRU_{HD}⁵) constitue notre simulation de référence pour l'apprentissage lors de la modélisation GAM.

4.4.1.2 La simulation à affiner

Nous avons choisi de descendre l'échelle de la simulation SGI (*Solar, Greenhouse Gas and Insolation*) [Servonnat, 2010] que nous nommons SGI_{LD}. Il s'agit d'une simulation mensuelle couvrant le dernier millénaire dans son intégralité de janvier de l'an 1000 à décembre 1999 sur 19 niveaux verticaux. C'est une simulation globale avec une grille de 96×72 , soit une résolution de 3.75° en longitude et 2.5° en latitude. Nous nommerons désormais cette simulation SGI_{LD}. Elle a été effectuée en appliquant des forçages solaires, orbitaux et de gaz à effet de serre. Le choix de cette simulation pour la modélisation des niches écoculturelles des Vikings est justifié dans la section 9.3.3.1.

Cette simulation possède les variables contenues dans le tableau 4.3. Les simulations SGI_{LD} et CRU_{HD} possèdent des variables en communs, mais afin de réaliser la régression nous devons d'abord homogénéiser, sélectionner et effectuer diverses manipulations de ces données.

5. Afin de distinguer les différentes résolutions des simulations avant ou après la descente d'échelle, nous leur attribuons les suffixes HD pour *High Definition* et LD pour *Low Definition*.

Tableau 4.3 : Liste des variables de la simulation SGI (*Solar, Greenhouse Gas and Insolation*).

Nom	Description	Unité
x	Longitude	° (décimaux)
y	Latitude	° (décimaux)
ndayrain	Nombre de jours de pluvieux (>0.1mm)	jours
precip	Précipitations	kg.m ⁻²
q2M	Humidité spécifique à 2m	%
rhum	Humidité relative	%
slp	Pression au niveau de la mer	Pa
snow	Chutes de neige	kg.m ⁻²
t2m	Température moyenne de l'air à 2m	K
t2m_max	Température maximale de l'air à 2m	K
t2m_min	Température minimale de l'air à 2m	K
vtu	Vent zonal (ouest à est)	m.s ⁻¹
vity	Vent méridional (sud à nord)	m.s ⁻¹

4.4.1.3 Traitement des variables avant le GAM

Avant d'établir le modèle GAM nous devons veiller à ce que les variables soient les mêmes dans la simulation de référence et dans la simulation dont on cherche à descendre l'échelle.

Homogénéisation des variables de CRU et de SGI

Nous avons renommé les variables de SGI_{LD} avec les noms de celles utilisées pour CRU_{HD} et converti les valeurs de ces dernières lorsque les unités étaient différentes. Pour CRU_{HD}, nous avons recréé les t_n et t_x à partir de la variable dtr (variation des températures journalières) et t_{mp} (températures moyennes) selon la formule équation 4.1 et équation 4.2, comme indiqué dans le fichier *readme.txt* associé à la simulation (<http://www.cru.uea.ac.uk/cru/data/hrg/tmc/readme.txt>, [New *et al.*, 2002])

$$t_n = t_{mp} - (0.5 * dtr) \quad (4.1)$$

$$t_x = t_{mp} + (0.5 * dtr) \quad (4.2)$$

Enfin nous avons supprimé les variables qui n'étaient pas en commun aux deux simulations.

Découpage temporel et spatial

SGI_{LD} possède des données pour chacun des mois de l'année sur 1000 ans, nous avons, à l'aide de commandes CDO et NCO⁶, découpé la simulation selon les périodes concernées puis fait la moyenne pour chacun des mois obtenant ainsi une simulation mensuelle pour

6. Logiciels sous licence GNU GPL permettant la manipulation de fichiers *NetCDF*, format dans lequel se trouve les simulations.

chaque période. Parallèlement, nous avons calculé les moyennes mensuelles de la simulation de référence CRU. Ensuite, pour l'ensemble des simulations nous avons procédé à un découpage spatial correspondant à la zone étudiée qui balaye les régions du Nord-Est du Canada au Nord-Ouest de la Grande Bretagne entre les longitudes 1.25°W et 85.75°W, et les latitudes 48.75°N et 85.25°N.

Une fois les données préparées de la sorte nous procédons à la construction du modèle GAM pour prédire les tn , tx et pr mensuelles. Ces trois variables constituent les prédicteurs de la modélisation GAM.

4.4.1.4 Description du processus de descente d'échelle par statistique GAM

La figure 4.1 liste les principales étapes suivies pour acquérir les températures extrêmes et les précipitations à haute résolution pour chaque mois de l'année. La procédure est composée de 3 étapes.

Apprentissage : élaboration de règles de régressions à partir de la simulation à haute résolution actuelle CRU_{HD} .

Validation des règles établies lors de l'apprentissage. À cette étape nous réalisons une prédiction GAM sur la simulation dégradée CRU_{LD} puis la comparons avec CRU_{HD} . Les règles de régressions sont validées si la prédiction est identique à la simulation HD.

Application des règles précédemment établies dans le but d'augmenter la qualité de l'échelle de la simulation SGI_{LD} , correspond à la phase de prédiction.

4.4.1.5 Dégradation de la simulation HD de référence

L'objectif consiste à dégrader volontairement la résolution d'une simulation climatique à haute résolution (CRU_{HD}) afin de voir si nous sommes en mesure, grâce à la régression GAM, de rétablir à partir de cette simulation dégradée (CRU_{LD}) une simulation HD relativement similaire à CRU_{HD} (partie apprentissage de la figure 4.1). Pour cela, nous avons procédé à la dégradation de CRU_{HD} de manière à obtenir CRU_{LD} à la même résolution que la simulation dont on cherche à augmenter la résolution (SGI_{LD} : $3.75^\circ \times 2.5^\circ$). Nous avons réalisé CRU_{LD} par agrégation de pixels, c'est-à-dire en affectant au pixel central la moyenne des pixels qui l'entourent.

4.4.1.6 Construction des prédicteurs

Les prédicteurs sont de plusieurs types : géographiques, topographiques et climatiques. Selon le type de prédicteurs, le degré de précision est plus ou moins élevé, en effet on peut obtenir aisément des prédicteurs géographiques et topographiques à des résolutions nettement plus élevées que les prédicteurs climatiques. Ainsi, les longitudes et latitudes ont un degré de précision ultime, et si l'on considère la topographie comme inchangée entre les

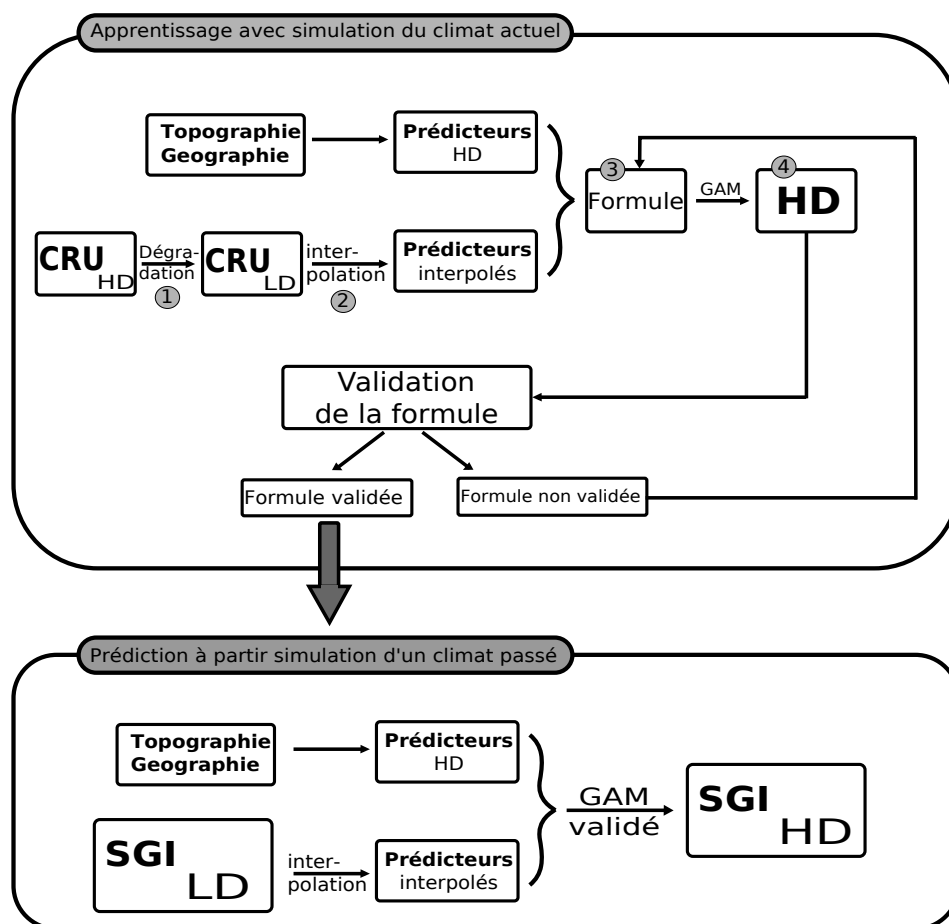


FIGURE 4.1 : Méthode de descente d'échelle statistique GAM. L'apprentissage est effectué sur une simulation à haute résolution actuelle (CRU_{HD}) que l'on dégrade (1) afin d'obtenir une simulation basse résolution (CRU_{LD}) qui nous procure les prédicteurs climatiques. On établit ensuite une formule qui combine prédicteurs climatiques, prédicteurs géographiques et topographiques (3) permettant la régression GAM dont l'application procure une simulation à haute résolution (4). On teste alors si la simulation obtenue équivaut à la simulation d'origine (CRU_{HD}) ; si la différence est significative alors on ré-établit une autre formule sinon on applique la formule validée à la simulation basse résolution dont on veut descendre l'échelle (SGI_{LD}) pour obtenir une simulation haute résolution (SGI_{HD}).

époques concernées et l'actuel, nous pouvons acquérir des prédictors, comme l'altitude, avec une résolution très élevée et une certitude accrue. En revanche, nous construisons les prédictors climatiques, car ils fluctuent dans le temps, en effectuant une interpolation spatiale bilinéaire des simulations LD (voir figure 4.1). Ainsi, la modélisation GAM mêle des prédictors de haute fiabilité (comme les coordonnées géographiques ou l'altitude) avec d'autres moins précis (comme l'humidité relative ou les températures moyennes). Prenant ces éléments en considération, la modélisation GAM doit être effectuée avec un choix pertinent de prédictors.

4.4.1.7 Choix des prédictors et élaboration de la formule

Le choix des prédictors varie en fonction des prédicteurs, on l'instaure à partir de relations présumées entre les variables qui, selon nous, expliquent au mieux le prédicteur (on sait par exemple que la température diminue en fonction de l'altitude et des longitudes). Pour chaque GAM effectué, on obtient un score de déviance expliquée lequel est d'autant plus élevé que la prédiction est correcte (tableau 4.4). Après avoir testé plusieurs combinaisons de prédictors, nous ne conservons que celles qui obtiennent les meilleurs scores pour cette statistique. Le tableau 4.5 indique quelles sont les combinaisons de prédictors retenues ainsi que les « formules » des modèles GAM indiquant les relations entre prédictors et prédicteurs. Ces relations peuvent être visualisées graphiquement figures 4.2, 4.3 et 4.4.

Chacune des formules commence par le terme désignant le prédicteur puis est suivie d'une somme indiquant les relations avec chacun des prédictors sélectionnés. Ces relations sont symbolisées par « $\sim s()$ ». On remarque notamment la présence de relations multiples permettant des lissages isotropiques comme $tn \sim s(x, y)$ ou $pr \sim s(reh, tmp)$. La première relation, établie en fonction de la position géographique, fournit des informations comme les influences de l'éloignement de la côte, de la calotte ou de tout élément massif du paysage sur le prédicteur. La seconde relation reflète l'effet de la conductance thermique induite par l'humidité sur les précipitations.

4.4.1.8 Validation de l'apprentissage et prédiction

Nous employons deux méthodes pour vérifier la cohérence de la formule GAM établie précédemment avant de l'utiliser pour affiner la résolution de SGI_{LD} . Premièrement nous testons statistiquement la différence entre la simulation d'origine (CRU_{HD}) et la simulation HD obtenue par modélisation GAM. Secondement nous procédons à une vérification visuelle afin de s'assurer que d'éventuelles variations importantes ne se situent pas dans les zones d'intérêts de l'étude⁷.

Les variances expliquées du modèle GAM retenu sont supérieures à 98% (scores très élevés) mais afin de tester statistiquement la cohérence de l'apprentissage, nous effectuons une régression entre le modèle obtenu à partir de CRU_{LD} et la simulation d'origine CRU_{HD} .

7. C'est-à-dire toutes les zones accessibles à la population étudiée.

Tableau 4.4 : Validation statistique de la modélisation GAM pour chaque prédictant (tn, tx, pr). La première colonne (Déviance), exprime la déviance expliquée par le modèle GAM et les autres colonnes correspondent aux résultats de la régression (de type $y = ax$) entre la simulation HD obtenue par la modélisation GAM et la simulation CRU_{HD} d'origine. On trouve le coefficient de régression (R^2), le coefficient directeur de la droite de régression (a), et la probabilité de non échec (p.value).

	Déviance	r2	a	p.value		Déviance	r2	a	p.value
Janvier	99.754	0.999	0.999	$<2.2.10^{-16}$	Janvier	99.868	0.999	0.999	$<2.2.10^{-16}$
Mars	99.891	0.999	0.999	$<2.2.10^{-16}$	Mars	99.952	0.999	0.998	$<2.2.10^{-16}$
Avril	99.914	0.999	0.998	$<2.2.10^{-16}$	Avril	99.96	0.998	0.997	$<2.2.10^{-16}$
Mai	99.938	0.997	0.995	$<2.2.10^{-16}$	Mai	99.94	0.993	0.993	$<2.2.10^{-16}$
Juin	99.937	0.994	0.99	$<2.2.10^{-16}$	Juin	99.938	0.991	0.992	$<2.2.10^{-16}$
Juillet	99.916	0.992	0.986	$<2.2.10^{-16}$	Juillet	99.935	0.994	0.993	$<2.2.10^{-16}$
Août	99.944	0.993	0.987	$<2.2.10^{-16}$	Août	99.958	0.993	0.992	$<2.2.10^{-16}$
Septembre	99.964	0.997	0.994	$<2.2.10^{-16}$	Septembre	99.984	0.994	0.992	$<2.2.10^{-16}$
Octobre	99.963	0.998	0.997	$<2.2.10^{-16}$	Octobre	99.986	0.997	0.996	$<2.2.10^{-16}$
Novembre	99.927	0.999	0.998	$<2.2.10^{-16}$	Novembre	99.962	0.998	0.998	$<2.2.10^{-16}$
Décembre	99.809	0.999	0.999	$<2.2.10^{-16}$	Décembre	99.889	0.998	0.999	$<2.2.10^{-16}$

(a) tn

	Déviance	r2	a	p.value
Janvier	98.707	0.973	0.965	$<2.2.10^{-16}$
Février	98.81	0.976	0.967	$<2.2.10^{-16}$
Mars	98.753	0.97	0.967	$<2.2.10^{-16}$
Avril	98.679	0.982	0.98	$<2.2.10^{-16}$
Mai	99.134	0.989	0.984	$<2.2.10^{-16}$
Juin	99.428	0.979	0.985	$<2.2.10^{-16}$
Juillet	99.659	0.992	0.996	$<2.2.10^{-16}$
Août	99.494	0.958	0.977	$<2.2.10^{-16}$
Septembre	99.222	0.932	0.964	$<2.2.10^{-16}$
Octobre	99.071	0.974	0.973	$<2.2.10^{-16}$
Novembre	99.101	0.967	0.968	$<2.2.10^{-16}$
Décembre	99.054	0.975	0.972	$<2.2.10^{-16}$

(b) tx

(c) pr

Tableau 4.5 : Prédicteurs et formules d'interpolation GAM des précipitations et des températures extrêmes.

Prédictant	Prédicteurs et formule employés
Température minimale (tn)	Altitude, longitude, latitude, précipitations, humidité relative $tn \leftarrow gam(tn \sim s(x,y) + s(y) + s(elv) + s(pr) + s(reh))$
Température maximale (tx)	Altitude, longitude, latitude, précipitations, humidité relative $tx \leftarrow gam(tn \sim s(x,y) + s(y) + s(elv) + s(pr) + s(reh))$
Précipitations (pr)	Altitude, longitude, latitude, températures moyennes, humidité relative $pre \leftarrow gam(pr \sim s(x,y) + s(y) + s(tmp) + s(reh) + s(elv) + s(reh,tmp))$

La droite de régression est de type $y = ax + b$, nous forçons cette droite à passer par l'origine (soit $b = 0$), ainsi si nous obtenons un coefficient directeur (a) proche de 1, cela signifie une grande ressemblance entre chacun des membres ($y \simeq x$), dans ce cas la simulation issue du modèle GAM équivaut à la simulation d'origine.

Nous avons obtenu des coefficients directeurs (a) très proches de 1 pour chacun des modèles GAM retenus pour chaque mois, dans chaque cas la régression a été certifiée par un coefficient de régression (R^2) très proche de 1 et une *p.value* toujours très proche de 0 (voir tableau 4.4). Cependant le modèle n'est pas parfait, nous pouvons observer comment se distribuent les résidus (c'est-à-dire les différences entre les valeurs modélisées et celles d'origine) dans la figure 4.5.

Nous nous apercevons que les moyennes des différences sont très proches de 0 et que

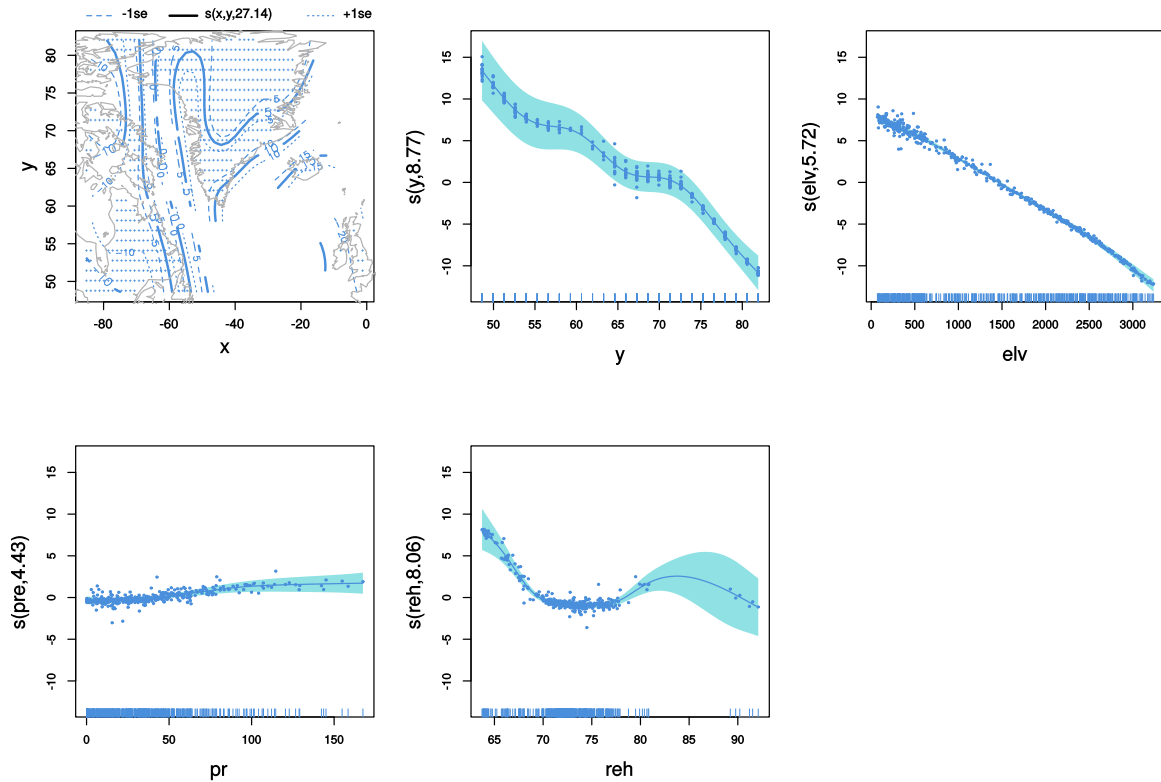


FIGURE 4.2 : Courbes de régressions ayant servi à la descente d'échelle des températures minimales (tn) du mois de janvier. En abscisses sont figurées les variables explicatives et en ordonnées leur relation (degré de liberté) avec la variable expliquée (tn). Les variables explicatives sont : l'altitude (elv), la longitude (x), la latitude (y), les précipitations (pr) et l'humidité relative (reh). Le premier graphique indique les variations de tn en fonction des longitudes et latitudes.

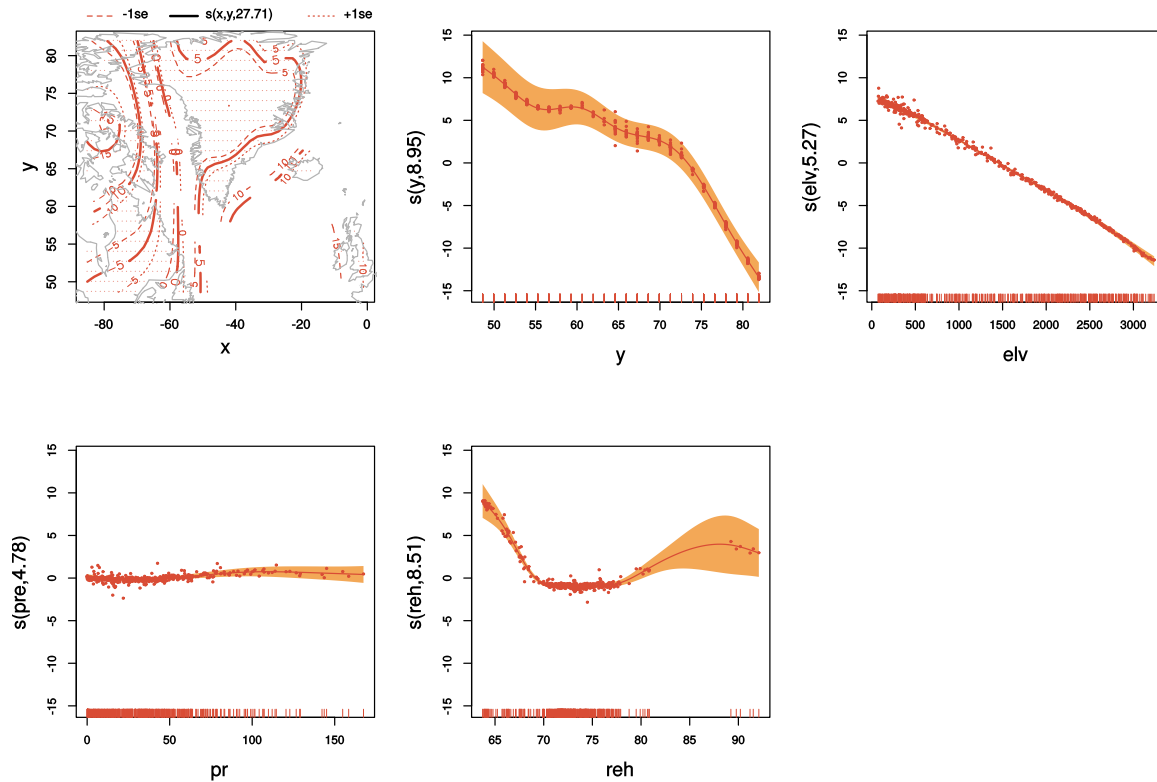


FIGURE 4.3 : Courbes de régressions ayant servies à la descente d'échelle des températures maximales (t_x) du mois de janvier. En abscisses sont figurées les variables explicatives et en ordonnées leur relation (degré de liberté) avec la variable expliquée (t_x). Les variables explicatives sont : l'altitude (elv), la longitude (x), la latitude (y), les précipitations (pr) et l'humidité relative (reh). Le premier graphique indique les variations de t_x en fonction des longitudes et latitudes.

la majorité des données reste groupée autour de la moyenne. Néanmoins, l'observation de nombreuses valeurs isolées nous conduit à mener une deuxième vérification qui est cette fois-ci visuelle. En effet, nous tolérons que la simulation modélisée présente des différences importantes avec CRU_{HD} , si et seulement si, elles se situent dans des zones où nous savons pertinemment que la population étudiée ne peut pas se trouver (dans le cas de la reconstruction de niches écoculturelles des Vikings, il s'agit par exemple de zones comme la calotte glaciaire). La figure 4.6 représente un exemple de cette vérification pour la température minimale (t_n) des mois de janvier et juillet. Nous constatons qu'il y a peu de variations entre ces cartes. Après avoir effectué ce contrôle pour chacun des prédictants et pour chacun des mois, nous sommes en mesure de valider les modèles GAM et pouvons

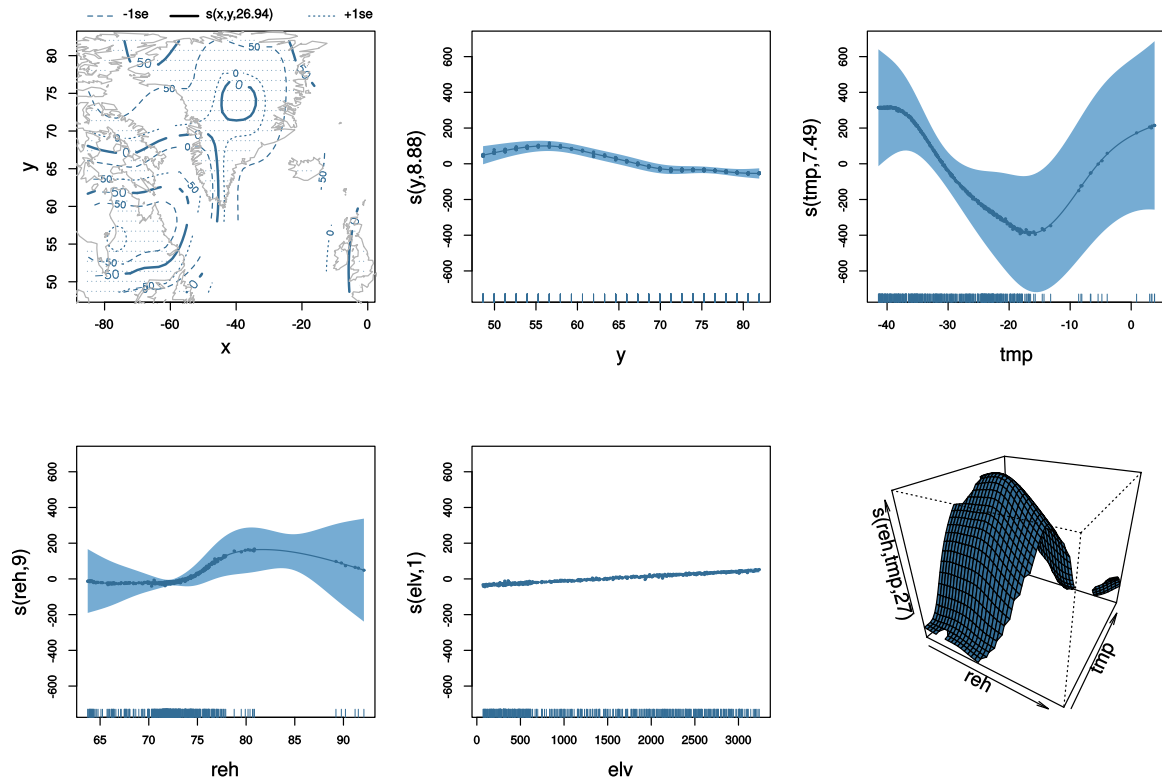


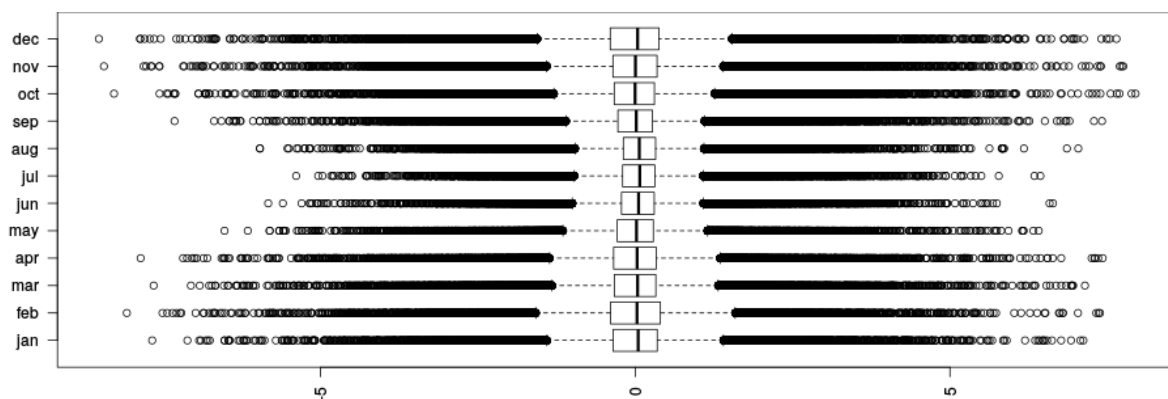
FIGURE 4.4 : Courbes de régressions ayant servi à la descente d'échelle des précipitations (pr) du mois de janvier. En abscisses sont figurées les variables explicatives et en ordonnées leur relation (degré de liberté) avec la variable expliquée (pr). Les variables explicatives sont : la latitude (y), la longitude (x), l'humidité relative (reh) et l'altitude (elv). Le premier graphique indique les variations de pr en fonction des longitudes et latitudes.

donc l'appliquer à la simulation SGI_{LD} afin d'obtenir les variables mensuelles désirées (tn , tx et pr) à haute résolution (partie prédiction de la figure 4.1).

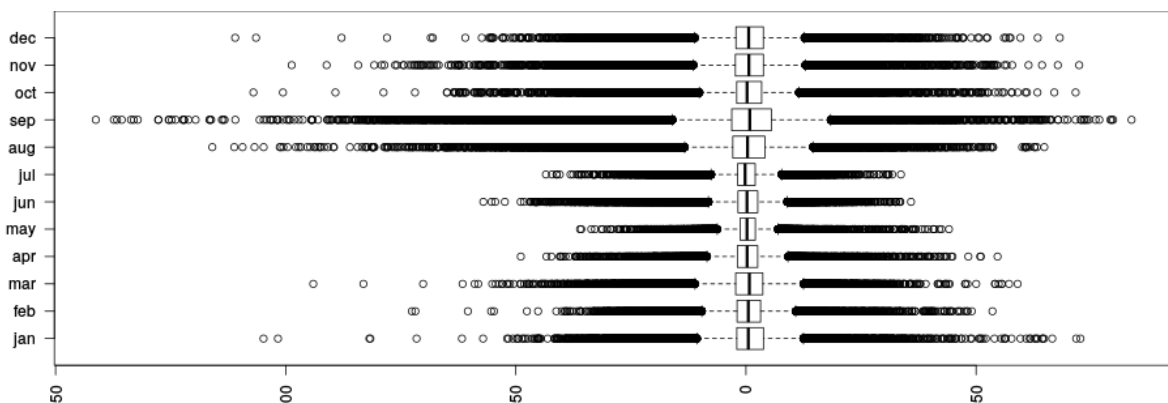
4.4.1.9 Calculer les variables bioclimatiques

Une fois tn , tx et pr obtenues à l'échelle désirée pour chacun des mois de l'année, nous avons recalculé les 19 variables bioclimatiques (tableau 2.6), en utilisant le *package* Dismo [Hijmans *et al.*, 2013] du logiciel R.

(a) Représentation entre les températures minimales (t_n) prédites et celles de la simulation d'origine



(b) Représentation entre les températures maximales (t_x) prédites et celles de la simulation d'origine



(c) Représentation entre les précipitations (pr) prédites et celles de la simulation d'origine

FIGURE 4.5 : Distribution des résidus provenant de la régression entre CRU_{HD} et la simulation à haute définition obtenue par modélisation GAM à partir de CRU_{LD}.

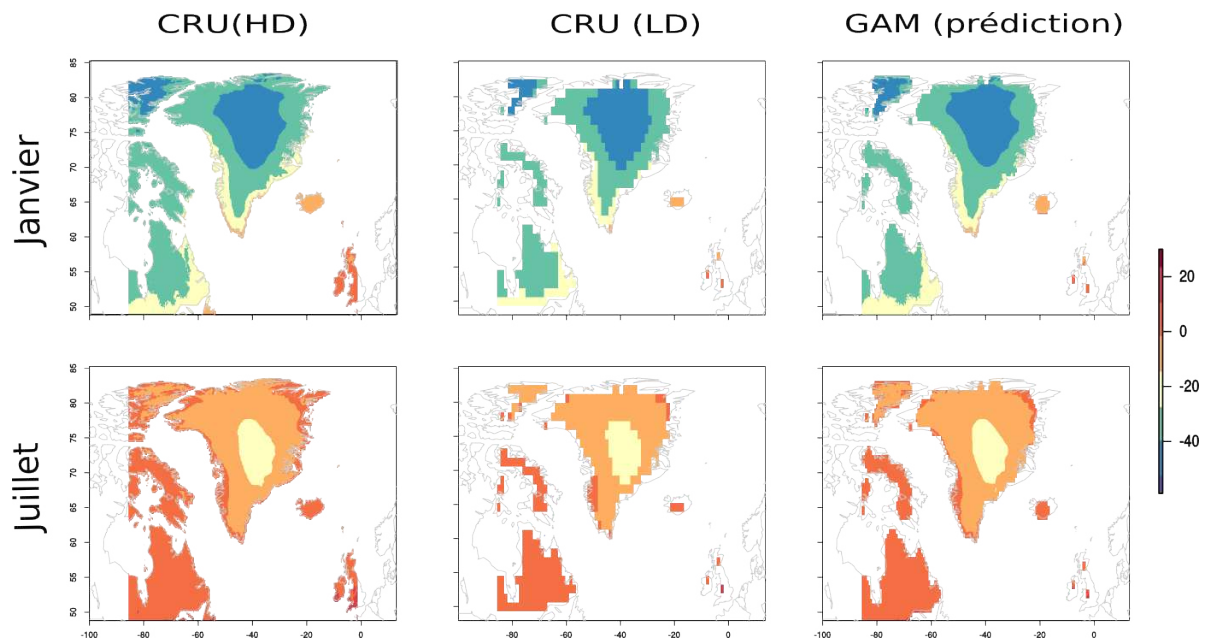


FIGURE 4.6 : Vérification visuelle de la descente d'échelle par régression GAM des températures minimales, exemple pour les températures minimales des mois de janvier (en haut) et de juillet (en bas). À gauche se trouvent les températures du modèle à haute définition (CRU_{HD}) utilisé pour le calibrage, au milieu le modèle basse résolution (CRU_{LD}) dont nous avons dégradé la résolution et à droite le modèle à haute résolution prédit par régression à partir du modèle basse définition. On observe que les modèles prédits sont très similaires aux modèles d'origine particulièrement dans la zone d'intérêt (Sud du Groenland). Cette comparaison a été répétée pour les douze mois de l'année pour chacune des trois variables, ceci nous permet de valider l'emploi de cette méthode de descente d'échelle. (Remarque : la détérioration de l'échelle suivie de l'interpolation crée, du fait du manque de données pour l'océan de CRU, un rognage au niveau de la côte. Un tel rognage n'a pas lieu pour la simulation SGI car elle n'a pas de masque au niveau de la mer).

4.5 Critiques de la méthode présentée

La méthode présentée reste encore perfectible, il faudrait pouvoir notamment tester de façon plus robuste la validité des résultats obtenus. Il serait aussi intéressant de la tester pour d'autres régions et d'autres époques. Enfin, intégrer d'autres modèles prédictifs que le GAM pourrait aussi augmenter significativement la fiabilité des résultats (par exemple des méthodes bayésiennes ou des réseaux de neurones). Cependant, les résultats obtenus, restent convenables et apportent la possibilité d'étudier finement les variations des niches au cours du temps et ainsi d'en apprécier leur dynamique évolutive.

Synthèse du chapitre : Obtenir des couches environnementales à haute définition pour une période spécifique

La qualité d'une prédiction de niche écologique ou écoculturelle est fortement dépendante de la résolution des couches de données environnementales utilisées pendant la modélisation. Il existe pour l'actuel ainsi que pour certaines époques passées et futures de telles données. Cependant, les modélisateurs se trouvent souvent freinés dans leur élan quand il s'agit de reconstruire une niche d'une espèce ou d'une culture dans un environnement pour lequel une simulation climatique à très haute résolution n'est pas disponible.

Nous présentons ici une méthode qui permet d'obtenir à partir de simulations climatiques globales basse résolution des couches de données climatiques à très haute résolution pour une région donnée et une époque couverte par la simulation.

Cette méthode permet d'obtenir des prédictions de niches pour des environnements ciblés. De plus, elle offre la possibilité de prédire ou de projeter des niches dans de nombreux environnements et d'apprécier ainsi la dynamique temporelle d'une espèce ou d'une culture.

Automatisation de la modélisation de niches

Préambule

Modéliser une niche écologique d'une population en utilisant un algorithme prédictif est chose relativement aisée et rapide grâce aux interfaces graphiques de type Desktop-GARP, mais si on multiplie le nombre d'espèces, le nombre d'algorithmes et le nombre de projections pour différentes périodes, on obtient rapidement une quantité importante de modélisations à réaliser. De plus, l'emploi des logiciels devient fastidieux et source d'erreurs dans ce cas. Il est donc utile dans un souci de commodité, mais aussi de répliquabilité facilitée et dépourvue d'erreurs, d'automatiser ces tâches. Nous allons voir dans ce chapitre comment l'emploi d'Open Modeller en mode console permet l'exécution simultanée de plusieurs algorithmes prédictifs ainsi qu'une série de tests sur les modèles obtenus.

Sommaire

5.1	Intérêt de multiplier les modèles	105
5.2	État de l'art	106
5.3	Scripts de génération de commande <i>om_console</i>	106

5.1 Intérêt de multiplier les modèles

L'étude d'une seule population dans un milieu est au final peu informative, en effet il est plus pertinent de comparer son écologie avec celle d'une autre espèce comme une compétitrice, ou d'autres espèces de son écosystème. Par ailleurs une étude de biodiversité portant

sur la modélisation de plusieurs espèces peut dévoiler des interactions non soupçonnées, comme l'identification d'une espèce clé de voûte dont la protection pourrait être cruciale.

L'intérêt de multiplier les modèles consiste aussi à augmenter l'exhaustivité d'une étude. Ainsi, l'on peut souhaiter augmenter le nombre d'espèces, comme énoncé précédemment mais aussi effectuer l'ENM en utilisant plusieurs algorithmes, utiliser des séries distinctes de variables environnementales ou varier les aires géographiques. Les modélisateurs peuvent également vouloir projeter les niches prédites dans des paléoenvironnements ou des environnements futurs. Dans un autre contexte, la multiplication des modèles peut servir à tester la pertinence de ces derniers.

Dans les cas sus-cités, le modélisateur se trouve très rapidement avec des dizaines voire des centaines de modèles à réaliser, cette entreprise est fastidieuse et le risque d'erreurs est d'autant plus grand que le nombre de tâches à effectuer est important.

5.2 État de l'art

Les interfaces utilisateurs dites *user friendly* n'existent pas pour la plupart des algorithmes prédictifs. Généralement leur emploi nécessite le lancement de scripts (codés en java, python, en langage C ou C⁺⁺), ou bien d'utiliser des *packages* ou des *plugins* d'autres logiciels comme R ou Matlab.

Des interfaces graphiques, comme MaxEnt, DesktopGARP ou *Open Modeller* permettent de modéliser les niches de plusieurs espèces à la fois. Ces logiciels permettent également de projeter ces dernières dans un environnement supplémentaire (notons qu'il n'est pas possible de projeter dans plusieurs environnements).

Certaines interfaces graphiques permettent, comme *Open Modeller*, un emploi simultané de plusieurs algorithmes sur les mêmes jeux de données initiaux (environnements et occurrences). Ceci évite à l'utilisateur de réitérer l'expérience autant de fois qu'il y a d'algorithmes et autant de fois qu'il y a d'espèces. En revanche, si l'utilisateur souhaite effectuer des projections dans *n* environnements supplémentaires, il devra alors recommencer *n* fois l'expérience.

Open Modeller et MaxEnt peuvent être lancés en mode console. Ceci leur confère un avantage particulier qui peut susciter l'intérêt des modélisateurs de niches. En effet, cette fonctionnalité offre la possibilité de grouper toutes les instructions pour l'ensemble des modélisations dans un seul script *batch* pouvant être ensuite exécuté en une seule opération.

5.3 Scripts de génération de commande *om_console*

Open Modeller possède un mode console assez élaboré avec notamment les commandes « *om_console* » et « *om_test* ». La première permet de lancer la modélisation à partir d'un fichier de requêtes conçu au préalable alors que la seconde permet l'exécution de tests,

comme l'*accuracy*, l'AUC ou le *partial ROC ratio*, à l'issue de la modélisation. Un fichier de requêtes contient toutes les informations nécessaires à la modélisation : chemins d'accès vers les occurrences, chemins vers les couches environnementales ainsi que l'algorithme à utiliser et ses paramètres. Ce fichier se présente sous la forme d'un fichier texte (.txt). Pour les besoins de cette thèse, nous avons généré automatiquement ces fichiers à l'aide d'un script R.

✍ Dans le cadre de notre exemple récurrent, nous souhaitons modéliser les niches écologiques de quatre espèces de lemmings à partir d'occurrences fossiles et actuelles (soit huit populations), pour deux environnements (l'actuel et le DMG), à l'aide de six algorithmes. Nous souhaitons également projeter chacune des niches prédites pour un environnement dans l'autre environnement. Nous devons donc effectuer un total de = 96 modèles (8 x 2 x 6). Il était par conséquent inenvisageable d'utiliser une interface graphique, ni même de saisir manuellement dans une console ces instructions. Nous avons, par conséquent, développé une fonction R permettant de générer des fichiers de requêtes pour chaque modèle ainsi qu'une commande « om_console » pour exécuter les requêtes et une commande « om_test » permettant de calculer les *partial ROC*. Finalement, nous avons appelé cette fonction en boucle pour chaque projection, chaque population et chaque algorithme. Les commandes générées ont été regroupées dans un fichier *batch* afin de lancer les instructions par lot.

Synthèse du chapitre : Automatisation de la modélisation de niches

Automatiser la procédure de modélisation permet de multiplier le nombre d'espèces modélisées dans de multiples environnements à l'aide de plusieurs algorithmes. La méthode présentée ici offre de nouvelles possibilités. Elle permet par exemple d'apprécier les variations de niches écoculturelles au cours du temps (chapitre 9) ou de discuter de diversité linguistique en fonction de l'environnement (chapitre 8).

Création de consensus de modèles de niches

Préambule

Le caractère stochastique des prédictions selon le paramétrage des algorithmes et la nature des données utilisées rendent impossible l'adoption d'une technique universelle pour la modélisation de niches. Ce chapitre présente plusieurs méthodes aidant à pallier les incertitudes de pertinence des prédictions qui consistent à sélectionner ou à réunir les prédictions issues de différentes techniques de modélisation.

Sommaire

6.1	Multiplier les techniques de modélisation	109
6.2	Présenter les résultats issus de plusieurs techniques de modélisation . . .	110
6.3	Sélection ou consensus : que choisir ?	116
6.4	Conception de consensus élitistes pondérés	117

6.1 Multiplier les techniques de modélisation

De nombreuses techniques de prédictions de niches ont été développées (chapitre 2). De manière à optimiser les résultats, les modélisateurs ont souvent recours à l'emploi simultané de plusieurs d'entre elles (par exemple [Araújo *et al.*, 2014, 2005; Roura-Pascual *et al.*, 2009; Thuiller *et al.*, 2013]). Or, comme vu dans la section 2.6, le caractère stochastique des algorithmes prédictifs peut conduire à des prédictions très différentes pour les mêmes données. Multiplier les modèles permet donc de confronter différentes prédictions de niches et optimiser les résultats. En revanche, une multitude de niches modélisées via l'utilisation de

nombreux algorithmes prédictifs permettra d'obtenir des modélisations plus performantes mais complexifiera la présentation et l'interprétation des résultats.

Il n'existe pas de technique universelle dont la fiabilité serait supérieure à celles de toutes les autres techniques [Qiao *et al.*, 2015]. Pour pallier ce manque, diverses stratégies permettant d'analyser les résultats sont déployées.

6.2 Présenter les résultats issus de plusieurs techniques de modélisation

Pour permettre une analyse fine et correcte, les modélisateurs se doivent de présenter et de synthétiser les résultats issus de plusieurs techniques de modélisations différentes. Cependant, une quantité importante de résultats très nuancés fait obstacle à cette synthèse. Dans un souci de clarté, les modélisateurs sélectionnent ou combinent les prédictions qu'ils obtiennent.

6.2.1 Sélection d'algorithmes

S'il n'existe pas de « balle d'argent », des algorithmes prédictifs de niches écologiques se démarquent en fournissant des prédictions dont l'exactitude est supérieure aux autres. MaxEnt et DesktopGARP (*Best subsets*) font preuve d'une assez grande notoriété parmi les modélisateurs et sont par conséquent les algorithmes prédictifs employés dans de nombreuses études (par exemple [Banks *et al.*, 2013a; Wang *et al.*, 2010]). Cette notoriété est principalement fondée sur des statistiques permettant d'évaluer la pertinence des modèles (par exemple [Allouche *et al.*, 2006]).

Les premières évaluations de modèles étaient basées sur des statistiques comme l'*accuracy* ou le kappa (section 2.6.3.3), mais rapidement de nombreux auteurs définissent l'AUC [Fielding et Bell, 1997] (section 2.6.3.5) comme critère de base pour retenir les modèles les plus pertinents [Hernandez *et al.*, 2006; Stockwell et Peterson, 2002]. Cependant nous devons garder toutes réserves quant à l'emploi exclusif de cette statistique pour l'évaluation de la capacité prédictive d'un modèle car des modèles distincts peuvent posséder la même valeur d'AUC (figure 6.1). Ainsi, il est préférable de multiplier les statistiques avant de juger de la pertinence d'un modèle. Des statistiques comme le *partial ROC ratio* [Peterson *et al.*, 2008a] (voir section 2.6.3.6), le TSS [Allouche *et al.*, 2006] (voir section 2.6.3.4), l'AIC (*Akaike information criterion*), le BIC (*Bayesian information criteria*) ou le LRT (*likelihood ratio tests*) (voir [Johnson et Omland, 2004 cités par Marmion *et al.*, 2009]) permettent également cette évaluation. Un autre test de pertinence consiste à écarter de la modélisation une partie des occurrences afin d'observer en quelle proportion ces dernières sont prédites par une nouvelle prédiction à partir des occurrences conservées. Certains algorithmes prédictifs, pour optimiser leur prédictions, emploient déjà ce dispositif (par exemple DesktopGARP *best subsets*, voir section 2.2.5.2).

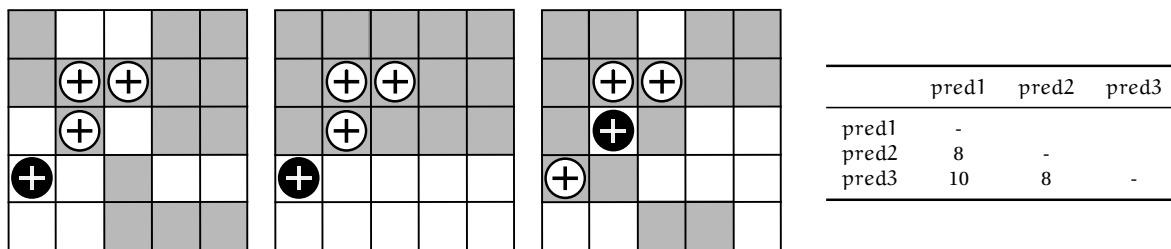


FIGURE 6.1 : Limite de la comparaison de pertinence des prédictions. Les grilles représentent trois prédictions de niches écologiques (de gauche à droite : pred1, pred2 et pred3) d'une espèce X obtenues à partir de trois algorithmes prédictifs différents. Les croix représentent les occurrences de présence, les cercles blancs les vrais positifs ; les cercles noirs les faux négatifs ; les cases grises les prédictions de présence ; et les cases blanches les prédictions d'absence. À droite est représentée la matrice des distances entre les prédictions (les valeurs indiquées correspondent au nombre de différences entre les prédictions).

Qiao *et al.* [2015] énoncent que les modèles les plus appropriés varient en fonction des organismes modélisés. Ainsi, ils préconisent de multiplier les modélisations en employant diverses architectures et d'analyser seulement les résultats issus des modélisations les plus performantes. La sélection des prédictions jugées les plus performantes permet de présenter les résultats sous formes intelligibles et facilite la discussion de ces derniers, cependant cette démarche comporte des biais pouvant s'avérer lourds de conséquences.

6.2.1.1 Les biais de la sélection

Lorsque des tests de performance basés sur la matrice de confusion (section 2.6.1) confèrent de fortes valeurs à plusieurs prédictions distinctes pour la modélisation de la même espèce, considérer chacune de ces dernières comme « vraie » est paradoxal puisqu'il n'existe qu'une seule « vraie solution » : la niche réelle. De plus, les probabilités sont très grandes pour que la niche réelle ne soit pas incluse parmi les solutions proposées car il est peu plausible que la niche réelle soit exactement représentée par une des différentes prédictions établies. La sélection d'une ou de quelques prédictions entraînent donc indubitablement de commettre des erreurs de commission et d'omission.

En multipliant les modèles, on peut être confronté à la situation théorique de la figure 6.1 où l'on observe trois prédictions distinctes qui comptent jusqu'à 10 différences sur 25 possibilités, mais qui possèdent pourtant les mêmes erreurs d'omission et de commission. Dans ce cas, il est *a posteriori* impossible de juger quelle est la prédiction la plus pertinente avec des statistiques basées sur la matrice de confusion. Une solution serait, comme indiqué précédemment, de procéder à des tests qui consisteraient à réitérer de nombreuses fois la modélisation avec une partie des occurrences retirées aléatoirement. Toutefois l'efficacité de cette méthode est dépendante du nombre d'occurrences ainsi que de la variabilité des conditions environnementales représentées par ces dernières (soit de leur distribution au

sein de la niche écologique). En archéologie par exemple, où les données peuvent être rares, ce test de pertinence peut être inadapté. Lorsque les statistiques ne permettent pas de sélectionner les prédictions de niches issues de différentes techniques de modélisation, alors seul un expert, connaissant l'écologie de l'espèce ou de la culture étudiée, est à même de sélectionner une prédiction. En outre, s'il s'agit de la modélisation d'une espèce actuelle, il peut être possible de vérifier, ou du moins d'apprécier, la cohérence des distributions prédites, mais l'exercice devient plus délicat lorsqu'on traite des données paléontologiques ou culturelles du passé dont la localisation des vestiges connues au jour de l'étude ne reflètent pas nécessairement la distribution réelle (biais taphonomiques, épistémologiques ou avancement de la recherche). La carte de l'expert doit donc être jouée avec prudence. De plus, cette dernière n'est pas dénuée de subjectivité, ce qui peut être un biais conséquent.

Comme l'on ne connaît pas la « formule » de la niche *a priori*, l'identification du meilleur outil pour la calculer n'est pas triviale. Il s'avère que choisir entre plusieurs distributions statistiquement significatives est objectivement impossible. Dans ce cas, même si une des prédictions est réellement la plus proche de la niche réelle, cette dernière est indécidable. Afin d'obtenir une solution très proche de la niche réelle, il est alors préférable de prendre en considération un maximum de solutions en élaborant un consensus.

6.2.2 Prédictions consensuelles

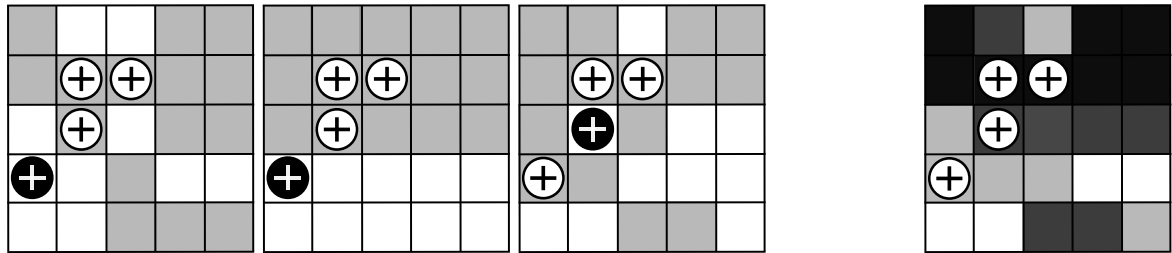
Les consensus sont des combinaisons de prédictions provenant de techniques différentes de modélisation, ils ont d'abord été employés en économie, en biomédecine, en météorologie, climatologie et hydrologie [Marmion *et al.*, 2009]. Le consensus présente l'intérêt de fournir une seule prédiction qui synthétise l'ensemble des prédictions faites par les divers algorithmes employés pour l'ENM ou l'ECNM.

6.2.2.1 Principaux consensus

Il existe plusieurs classes de consensus, les principales sont les suivantes :

- les « consensus simples », construits à partir de la somme de chaque prédiction ;
- les « consensus moyens » qui considèrent toutes les prédictions en faisant la moyenne de celles-ci ;
- les « consensus médians » qui proposent une solution en accord avec au moins la moitié des prédictions ;
- les « consensus élitistes », générés à partir d'une sélection d'algorithmes selon leur pertinence, laquelle peut être évaluée par l'AUC ;
- les « consensus pondérés » pour lesquels des poids sont attribués aux différentes prédictions (la pondération peut être affectée à toutes les autres classes).

Il est fréquent de transformer les prédictions issues des différentes architectures de modélisation en prédictions binaires¹ avant de construire un consensus. De cette façon, un consensus simple obtenu à partir de la somme de prédictions binaires (figure 6.2) produit une prédiction consensuelle dans laquelle les valeurs correspondent aux accords entre les prédictions. Un tel consensus peut être aussi désigné de « consensus fréquentiste ».



(a) Trois prédictions de niches écologiques d'une espèce X obtenues à partir de trois algorithmes prédictifs différents. ; cellules grises : présences (valeur 1) ; cellules blanches : absences (valeur nulle)

(b) Consensus, somme des trois modèles

FIGURE 6.2 : Construction d'un consensus simple ; croix : occurrences de présences ; ronds noirs : faux négatifs ; ronds blancs : vrais positifs. Dans le consensus (b) l'intensité du gris signale le nombre de prédictions en accord (gris clair : aucun accord ; gris foncé : deux accords ; noir : trois accords). Les cases blanches correspondent aux prédictions d'absence communes à tous les modèles.

6.2.2.2 Qualité d'un consensus

Dans la figure 6.2(b), le consensus est obtenu en additionnant les modèles binaires. Le consensus affiche une variation de probabilité de présence selon si la zone a été prédite par aucun, un, deux, ou l'ensemble des trois modèles. On remarque également que la surface de prédictions de présences (c'est-à-dire la représentation de la niche prédite) est supérieure à celle des autres modèles, ceci étant dû à une augmentation du nombre de faux positifs (augmentation de l'erreur de commission). À l'inverse, l'erreur d'omission diminue car le pourcentage de faux négatifs (prédictions d'absence où se trouve une occurrence de présence) baisse. Ces deux variations induisent une AUC différente de celle des autres modèles. Si l'AUC du consensus est supérieure à celle des modèles qui le compose, alors on estimera qu'il est mélioratif. Il est donc possible d'estimer la qualité d'un consensus en comptant le nombre de modèles qui ont une AUC inférieure à celle du consensus équation 6.1. Plus cette valeur est proche de "un" meilleur est le consensus, si cette valeur est médiocre, il est alors recommandé de présenter les résultats des modèles les plus performants.

$$\text{qualité}_{\text{consensus}} = \frac{\text{nb}_{\text{AUC}_{\text{modèle}} < \text{AUC}_{\text{consensus}}}}{\text{nb}_{\text{modèles}}} \quad (6.1)$$

1. Une prédiction binaire attribue une valeur de 0 pour l'absence et de 1 pour la présence.

6.2.2.3 Forces et faiblesses d'un consensus

Les consensus sont constitués pour pallier aux difficultés rencontrées quand plusieurs prédictions sont jugées pertinentes. Leur construction est basée sur l'hypothèse qu'une présence en un point prédite par un certain nombre d'algorithmes a plus de chance d'être vraie qu'une présence prédite par un nombre réduit d'algorithmes en un autre point. L'aspect cumulatif de la présentation consensuelle des résultats produit une accumulation des prédictions exactes comme des erreurs de l'ensemble des modèles qui le composent.

Puisqu'ils combinent différents modèles, les consensus gommement les effets de sur-apprentissage ou de sous-estimation. Cela leur confère un avantage par rapport aux modélisations effectuées à partir d'un seul algorithme prédictif. Regroupant les prédictions issues de plusieurs modélisations les consensus permettent de réduire le taux de faux négatifs, mais parallèlement le taux de faux positifs augmente (figure 6.2(b)). Une partie de ces faux positifs correspond réellement à la niche écologique de l'organisme étudié, mais une autre partie est une aberration des algorithmes. L'ensemble de ces prédictions abusives, résultant de la somme des prédictions de faibles probabilités de présences de plusieurs algorithmes est appelé « bruit de fond ». Ce dernier peut être très important si les prédictions n'ont pas été seuillées au préalable. Ainsi, en théorie le consensus simple d'une quantité infini de prédictions non seuillées présenterait une solution consensuelle dans laquelle une présence serait prédite en chaque point de grille. Un tel résultat serait alors dépourvu d'intérêt. Au vu de cela nous concluons, premièrement de l'importance d'appliquer un seuil aux prédictions avant de les regrouper et secondairement de l'utilisation d'un nombre de modèles raisonné, car si multiplier les prédictions augmente la performance prédictive du consensus, un nombre trop important peut nuire à sa qualité.

6.2.2.4 Les consensus dans *Open Modeller* et *Biomod*

Deux outils (le logiciel *Open Modeller* [De Souza Muñoz *et al.*, 2011] et le *package* R *Biomod* [Thuiller *et al.*, 2009]) permettent l'emploi simultané de plusieurs techniques de modélisations. Les deux structures offrent la possibilité de générer des centaines de prédictions car le premier possède un mode console et le second a des fonctions pouvant être appelées dans des scripts de programmation (voir chapitre 5).

Open Modeller possède une fonction pour créer des consensus, seulement cette fonctionnalité est limitée. En effet, cette fonction ne permet pas la réalisation de consensus de plus de cinq prédictions. De plus, le fait que les prédictions doivent être rendues binaires (suivant le LPT section 2.5.2) par le logiciel avant d'être combinées, provoque une perte de la variabilité des prédictions fournies par les modèles originels. Respectant ces contraintes, *Open Modeller* permet de créer des consensus simples, moyens ou médians. Une pondération est possible, mais l'affectation des valeurs ne peut être faite automatiquement. La création de consensus élitistes est possible mais la sélection doit se faire manuellement, opération qui peut devenir fastidieuse quand le nombre de modèles est conséquent.

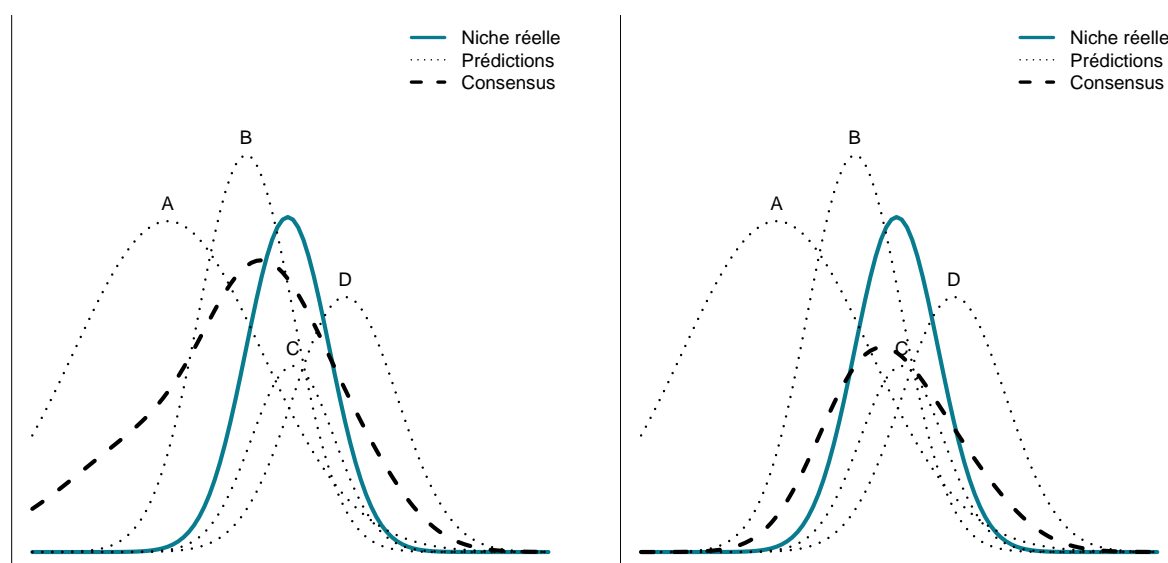
Récemment une mise à jour de Biomod (Biomod2 [Thuiller *et al.*, 2013], voir section 2.2.5.3) a été effectuée. Cette nouvelle version, dans laquelle a été implémentée des statistiques telles l'AUC, le TSS ou le Kappa, permet la création de consensus des différentes classes (voir section 6.2.2.1). Zhang *et al.* [2015] comparent une série de consensus fréquentistes, médians et moyens élitistes selon deux statistiques ($AUC > 0.7$ et $TSS > 0.4$) réalisés à partir de modèles créés avec Biomod. Cette publication montre à quel point ce *package* offre de réelles perspectives pour la création de consensus et l'évaluation de leurs performances.

6.2.3 Comparaison consensus classiques et consensus élitistes pondérés

Nous appelons consensus « classiques », les consensus simples et les consensus moyens. Nous avons vu que l'emploi de techniques diversifiées permettait d'obtenir des consensus plus robustes. Seulement, la diversification induit d'incorporer dans la modélisation des algorithmes performants ainsi que d'autres moins adaptés. Il est alors possible par la pondération de faire en sorte que le consensus reflète proportionnellement la pertinence des modèles. Les poids peuvent être attribués arbitrairement (selon la notoriété de l'algorithme) ou selon l'avis d'un expert au vu des projections géographiques de chacune des prédictions. Mais il est d'usage, car moins subjectif, d'effectuer la pondération selon des statistiques comme les scores d'une ACP [Thuiller, 2003], l'*accuracy*, la TSS ou l'AUC. La dernière étant largement plus utilisée que les autres à cet effet (par exemple [Poulos *et al.*, 2012]).

La figure 6.3 schématise une modélisation de niche théorique avec quatre algorithmes (A, B, C et D). Les prédictions de chacun des algorithmes sont représentées par des courbes en cloche à partir desquelles deux consensus ont été réalisés : un consensus moyen figure 6.3(a), obtenu en faisant la moyenne des courbes et un consensus pondéré figure 6.3(b) lequel a été obtenu en faisant une moyenne pondérée. La pondération a été effectuée arbitrairement selon la ressemblance des courbes avec la niche réelle. On note que trois prédictions (B, C et D) sont très proches de la niche réelle, elles sont donc affectées d'un poids élevé. La prédiction A, très médiocre, a quant à elle été exclue en lui affectant un poids « nul ». Le consensus moyen est moins proche de la niche réelle que ne l'est le consensus élitiste pondéré. Cette figure nous montre comment une prédiction médiocre incluse dans un consensus peut détériorer drastiquement sa qualité prédictive et donc l'intérêt de l'exclure de la prédiction consensuelle.

Afin de déterminer la solution la plus correcte de présenter les résultats entre le meilleur modèle ou les différents consensus : classiques, médians ou élitistes pondérés, Marmion *et al.* [2009] ont procédé à la modélisation de 28 espèces végétales. Les auteurs concluent que le consensus élitiste pondéré fournit les résultats les plus robustes. La deuxième meilleure méthode de consensus étant le consensus classique.



(a) Consensus moyen. La niche prédite par ce consensus qui est la moyenne des prédictions des algorithmes A, B, C et D (représentées en pointillés) est plus proche de la véritable niche que ne l'est chacune des prédictions de ces algorithmes prises séparément, nous remarquons toutefois que la prédiction du consensus est relativement large.

(b) Consensus élitiste pondéré. Les prédictions des algorithmes A, B, C et D ont été respectivement affectées des poids suivants : 0 (exclusion de la prédiction), 1.9, 1.8 et 1.7 puis la moyenne a été calculée. On obtient ainsi un consensus proche de la niche réelle.

FIGURE 6.3 : Consensus simple et consensus pondéré. La niche prédite par le consensus pondéré est plus fidèle à la niche réelle que ne l'est le consensus obtenu par simple addition (consensus simple). Remarque : une pondération de 0 permet d'écarter une prédiction, un poids inférieur à 1, de minimiser son importance alors qu'au contraire affecter un poids supérieur à 1 permet d'augmenter la représentativité de la niche prédite par cet algorithme. Les poids peuvent être définis selon des critères statistiques ou bien selon le jugement du modélisateur qui évalue, grâce à ses connaissances quant à l'écologie de l'espèce, la pertinence des prédictions.

6.3 Sélection ou consensus : que choisir ?

Si GARP et MaxEnt sont depuis longtemps considérés comme les algorithmes prédictifs les plus performants, d'autres techniques de modélisations aujourd'hui peuvent les surpasser pour des cas d'études donnés en des lieux donnés. Ainsi, il est impossible de connaître à l'avance quelle technique privilégier pour une modélisation particulière.

L'AUC ou le *partial ROC* permettent d'évaluer objectivement le potentiel de prédiction des algorithmes, cependant nous avons vu que l'on ne peut pas accorder une confiance aveugle à ces tests statistiques (6.1). Ainsi, sans connaissance *in vivo* de l'espèce ou de la culture étudiée, il est délicat de sélectionner un seul ou un petit nombre de prédictions issues de techniques de modélisations différentes. De plus, la quantité de prédictions

amplifie les difficultés d'interprétation *a fortiori* si ces dernières diffèrent grandement entre elles.

Lors de toute modélisation, le risque de ne pas utiliser l'outil adapté pour obtenir le résultat le plus fidèle à la réalité est existant. Libre à nous d'utiliser les techniques les plus variées pour ensuite les sélectionner. Toutefois, la sélection peut anéantir l'effort précédemment effectué, en écartant de la liste des prédictions retenues des prédictions localement ou globalement plus exactes que l'ensemble des prédictions sélectionnées. Pour cette raison, nous pensons qu'il est préférable d'utiliser des consensus malgré les défauts que ces derniers comportent comme le bruit de fond (section 6.2.2.3) mais qui peuvent être atténués par la pondération.

La création de consensus élitistes pondérés nous semblent être le meilleur compromis car elle permet de lisser les défauts du consensus et d'écarter les algorithmes les moins performants.

6.4 Conception de consensus élitistes pondérés

Dans le cadre de cette thèse nous avons utilisé le *mode_console* de *Open Modeller* afin de générer des prédictions ou des projections de niches en employant plusieurs algorithmes afin d'optimiser la modélisation (voir chapitre 5). Ainsi, pour une interprétation pertinente des résultats, nous avons opté pour la création de consensus élitistes pondérés.

L'automatisation des ECNM nous a permis pour chaque cas d'étude (chapitres 8,9 et 10) d'obtenir de nombreuses prédictions, mais avant de combiner les prédictions ou projections provenant des différents modèles, il faut s'assurer de leur homogénéité.

6.4.1 Rendre comparables les prédictions

Les prédictions de niches sont sous forme de couches d'informations géographiques comportant une valeur en chaque point de grille. Cette valeur correspond à la probabilité de présence estimée par le modèle. Or, des techniques de modélisations différentes impliquent des valeurs de prédictions propres à chacune. Par exemple, une prédiction de présence d'une valeur de 0.6 obtenue pour un point de grille avec l'algorithme MaxEnt a un sens différent qu'une prédiction de même valeur obtenue avec GARP. Sans un traitement préalable, il est donc impossible de comparer ces deux prédictions ou de les combiner dans un consensus. Il est donc nécessaire de procéder à une homogénéisation des prédictions avant de construire un consensus [Araújo et New, 2007].

L'homogénéisation des prédictions se réalise par l'application d'un seuil défini selon les valeurs prédites aux points d'occurrences (section 2.5.2). Le seuil le plus fréquemment employé est le LPT (*lowest presence threshold*).

Une fois seuillée, la prédiction est soit rendue binaire avec une valeur égale à 1 pour toutes les prédictions qui dépassent le seuil et 0 pour les prédictions inférieures, soit trans-

formées en prédictions variant entre 0 et 1 (où la valeur 0 est accordée aux prédictions égales au seuil et la valeur 1 à la valeur maximale de la prédiction). La transformation ou « remise à l'échelle » est préférable car elle permet de conserver la variabilité spécifique à chaque prédiction.

6.4.2 Exemple de création d'un consensus élitiste pondéré

✍ Nous reprenons notre exemple du lemming pour illustrer la procédure suivie afin de créer un consensus élitiste pondéré à partir de six algorithmes (figure 6.5). Les prédictions ont été effectuées sous *Open Modeller*, seuillées selon la méthode du LPT (*Low Prediction Threshold*) puis remises à l'échelle entre 0 et 1. Les diverses architectures de modélisations fournissent des prédictions différentes, rendant toute analyse des résultats impossible. Les statistiques de pertinence des algorithmes sont présentées tableau 6.4(b). Nous appliquons alors la méthode décrite précédemment et réalisons un consensus élitiste pondéré suivant les trois étapes décrites figure 6.4.







étape 1 La sélection des prédictions s'effectue selon une AUC supérieure à 0.8. Sont retenues les prédictions des algorithmes : GARP, SVM et *Environmental Distance*

étape 2 La pondération est effectuée selon l'évaluation de l'expert (après examen des projections géographiques de chaque prédiction) et selon le *partial ROC ratio*.

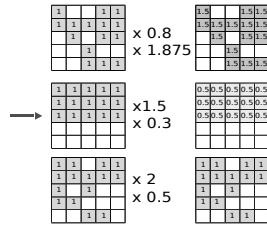
étape 3 La moyenne des prédictions pondérées est effectuée pour réaliser le consensus présenté figure 6.5(g).

$$w.\text{consensus} = \sum_i \text{algo}_i \times \text{Éval.expert}_i \times \text{pRocRatio}_i$$

(a) Formule d'un consensus pondéré : les algorithmes dont l'AUC est supérieure à une valeur fixée sont sélectionnés puis pondérés par l'évaluation donnée par l'expert et la valeur de leur *partial ROC ratio* respectif.

		AUC	Éval. expert	Roc ratio
GARP		0.89	0.80	1.875
MaxEnt		0.80	0.50	1.61
Env. Score		0.85	0.00*	1.70
SVM		0.81	0.30	1.50
RF		0.76	0.00*	1.02
Env. Dist		1.00	0.50	2.00

(b) **Étape 1** : sélection des meilleurs modèles. Seules sont retenues les prédictions avec une $AUC \geq 0.8$. (*) Notons que l'attribution d'une note de 0 par l'expert revient à écarter le modèle (l'évaluateur attribue des notes aux prédictions comprises entre 0 et 1 selon leur réalisme par rapport à l'écologie de l'espèce). En gras figurent les modèles retenus.



(c) **Étape 2** : pondération de chacun des modèles sélectionnés à l'étape 1 par l'évaluation donnée au modèle par l'expert et selon la valeur de leur *partial ROC ratio*, statistique performante afin d'évaluer l'exactitude d'une prédictions [Peterson *et al.*, 2008a].

3	1.5	0.5	3	3	
3	3	3	3	3	3.0
1.5	2	1.5	2	2	2.5
1	1	1.5	0	0	2.0
0	0	2.5	2.5	1.5	1.5
					0.5

(d) **Étape 3** : création du consensus en additionnant les modèles pondérés à l'étape précédente.

FIGURE 6.4 : Création d'un consensus pondéré à partir de plusieurs modèles dont le potentiel prédictif a été évalué statistiquement et par un expert. Ici, les différentes prédictions des algorithmes sont représentées de façon binaire (0 - 1) afin de simplifier la figure, mais il faut noter que la méthode proposée dans cette thèse permet de conserver la variabilité des prédictions au sein des algorithmes.

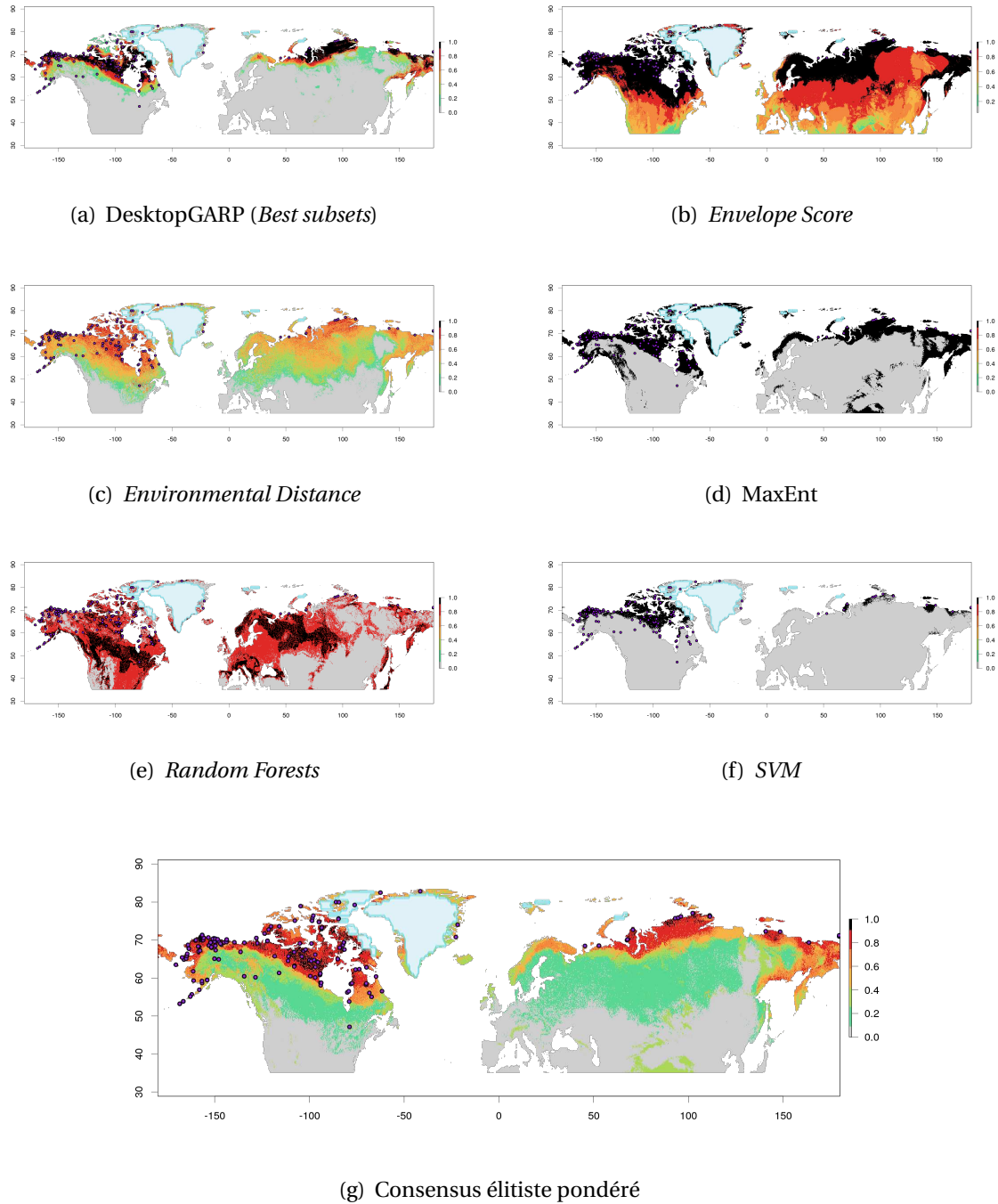


FIGURE 6.5 : Caractère stochastique des algorithmes prédictifs et consensus. Prédiction de la niche écologique de *Dicrostonyx sp* par différents algorithmes et consensus. Le consensus élitiste pondéré a été réalisé pour les prédictions avec une $AUC \geq 0.8$ et pondéré selon l'évaluation d'un expert et le *partial ROC ratio*.

Synthèse du chapitre : Création de consensus

L'emploi de techniques de modélisations diversifiées est préconisé pour l'obtention de résultats robustes. Cependant un grand nombre de modélisations peut conduire à des prédictions très nuancées. Un effectif élevé de prédictions peut nuire à l'interprétation des résultats. La sélection d'algorithmes selon des critères statistiques permet d'écarter les prédictions aberrantes, toutefois si deux prédictions distinctes sont évaluées très pertinentes, les départager relève de l'impossible. Une alternative est la création de consensus élitistes pondérés ; ces derniers écartent les modèles erronés et créent une prédiction consensuelle accordant un poids relatif à la performance des prédictions retenues.

Analyse des niches écologiques

Préambule

Ce chapitre a pour vocation de présenter les différents outils que nous employons pour analyser les prédictions de niches écoculturelles présentées dans la troisième partie de cette thèse. Il présente des questions auxquelles ces outils apportent des éléments de réponse et définit des termes qui nous serviront notamment à la comparaison de niches et à la description de leur évolution.

Sommaire

7.1	Regard critique sur les données	123
7.2	Examiner la répartition géographique des prédictions	124
7.3	Mesures et comparaison de niches	124
7.4	Identification et quantification des principaux facteurs environnementaux	125

7.1 Regard critique sur les données

Avant de se lancer dans l'analyse d'une prédiction de niche écoculturelle, nous devons être conscients des biais expérimentaux induits par les techniques de modélisation ou par les données elles-mêmes. Pour effectuer les ECNM de cette thèse, nous utilisons des consensus élitistes pondérés (section 6.4) ; ces derniers permettent d'optimiser les prédictions et dans une certaine mesure, de ne plus nous soucier de la fiabilité des modèles (voir chapitre 6). Les données, quant à elles, doivent être considérées avec plus de discernement. Nous avons vu dans les chapitres 1 et 2 que la nature et la résolution des variables environnementales pouvaient avoir un fort impact sur la modélisation. Au cours de ces chapitres, nous

avons également discuté de la représentativité des occurrences. Il faut donc s'assurer de l'homogénéité de l'effort d'échantillonnage ainsi que de la quantité suffisante d'occurrences (section 2.3.1.1). Enfin, la surface du M doit également être prise en considération car elle influe nettement sur l'issue de la prédiction (voir section 2.4.2.2). Effectivement, la mobilité des groupes humains peut varier d'une région à l'autre ou selon les époques, en fonction de leurs technologies ou de leur stratégies de subsistance par exemple [Binford, 1980; Kelly, 1983, 1992].

7.2 Examiner la répartition géographique des prédictions

Le premier réflexe des modélisateurs est de projeter la prédiction de la niche modélisée sur une carte pour pouvoir la visualiser géographiquement.

7.2.1 Comparaison avec nos connaissances sur la culture

L'observation de l'expression géographique d'une niche écoculturelle est en soit très informative. Effectivement, le modélisateur peut instantanément vérifier si la prédiction est en accord ou si elle va à l'encontre de théories établies en raisonnant sur l'ensemble de la prédiction ou en portant une attention particulière dans des régions à questionnements. Notons toutefois que ces observations bien que très informatives restent descriptives. Pour compléter ces analyses, l'expression géographique des prédictions peut être croisée à la répartition d'éléments de l'environnement.

7.2.2 Comparaison avec l'environnement

Les éventuelles similitudes entre une projection géographique et des cartes représentant des données écologiques ou biologiques comme la répartition d'une espèce, la durée de saison de croissance des végétaux, ou la répartition des biomes, peuvent révéler des traits écologiques. La confrontation des cartes de facteurs environnementaux tels que la température ou la topographie et des niches écoculturelles peuvent infirmer ou confirmer l'existence de barrières naturelles.

Ces comparaisons géographiques permettent également de soulever de nouvelles questions *a fortiori* lorsque l'on compare les prédictions de plusieurs cultures contemporaines ou les expressions géographiques d'une même culture à plusieurs époques.

7.3 Mesures et comparaison de niches

Plusieurs indices permettent d'évaluer la taille des niches ou d'estimer leurs ressemblances mutuelles. Bien qu'ils puissent individuellement être traduit en français, ce document de thèse utilisera les termes anglais pour ces mesures, afin que le lecteur puisse

aisément établir le rapprochement avec la littérature scientifique. Ces mesures sont les suivantes :

Breadth ou « largeur écologique », cette mesure reflète l’amplitude de l’ensemble des variations environnementales qui compose une niche.

Overlap ou mesure de « chevauchement de niches », sa valeur est proportionnelle à l’ensemble des lieux où les conditions environnementales sont propices aux conditions de vie des espèces concernées, en d’autres termes, cette mesure correspond aux lieux où les niches se recoupent.

Similarity ou similarité, cette mesure permet d’évaluer la ressemblance entre deux modèles. Sa valeur évolue entre 0 et 1. La valeur 1 indique deux modèles parfaitement identiques et la valeur 0 deux modèles entièrement distincts.

Ces mesures sont calculées à partir des indices d’Hellinger et de Schoener [Warren *et al.*, 2008] construits pour évaluer les champs de variations de l’environnement dans lesquels une espèce peut vivre [Carnes et Slade, 1982; Levins, 1968; Soberón, 2007; Soberón et Nakamura, 2009]. ENMTools [Warren *et al.*, 2010], SDMtools [VanDerWal *et al.*, 2011] et dismo [Hijmans *et al.*, 2013] sont des outils permettant d’effectuer ces mesures.

Le *package* R *phyloclim* [Heibl et Calenge, 2013] permet d’obtenir deux matrices triangulaires de similarité basées sur la distance D de Schoener [1968] pour la première et sur la distance de Hellinger [Van der Vaart, 1998] pour la seconde. La conversion de ces matrices en matrices de distance¹ permet la construction d’arbres. Nous nommons ces arbres obtenus à partir de niches écologiques : « arbres phyloécologiques », de tels arbres permettent d’apprécier les distances entre systèmes adaptatifs de groupes humains.

7.4 Identification et quantification des principaux facteurs environnementaux

Selon Hutchinson [1957], la niche écologique peut être vue comme un « hypervolume » (section 1.2.3.2) pour lequel chacune des dimensions correspond à une variable environnementale. La représentation de cet hypervolume par une analyse en composante principale (ACP) est donc particulièrement adaptée. Dresser l’ACP d’une prédiction de niche nous renseigne sur la participation des variables environnementales dans la composition de la niche. On peut ainsi déterminer si les populations étudiées sont soumises à des contraintes environnementales et auquel cas les identifier et les quantifier.

Les individus de l’ACP correspondent aux prédictions de présence et les variables de l’analyse aux données environnementales utilisées pour la modélisation. Bien que les prédictions aient été seuillées (section 2.5.2), le nombre de variables, la résolution de celles-ci et la surface couverte par la prédiction donnent lieu à une grande quantité d’informations.

1. la conversion consiste à retirer 1 à la matrice de similarité

Ainsi le tableau à réaliser pour l'ACP peut atteindre une taille colossale et au final produire une analyse faussée par les prédictions où la présence est prédite avec les plus faibles probabilités. Pour pallier le bruit de fond généré par ces prédictions, nous optons pour sélectionner celles-ci selon un seuil défini après observation de la projection géographique de la niche. Enfin, dans le souci de conserver la variabilité de la prédiction, nous pondérons les individus de l'ACP par les valeurs de probabilités de présence que leur avait attribuées le modèle.

Nous dénommons les ACP menées pour une prédiction donnée : « ACP individuelles » par opposition à d'autres analyses que nous menons qui combinent les prédictions de plusieurs niches dans une même ACP.

7.4.1 Les « ACP communes »

Dans l'objectif de comparer les niches écologiques entre elles, nous combinons dans une même ACP les prédictions de plusieurs niches. Ces prédictions sont sélectionnées suivant la même procédure que pour les ACP individuelles et sont également pondérées. Pour une meilleure lecture de l'analyse notre raisonnement s'appuie sur les ellipses de dispersion. Ces ellipses représentent la projection moyenne de l'inertie des individus de l'ACP.

Les intérêts de réaliser une telle analyse sont multiples. Considérant l'ACP produite comme le reflet de l'« hyperespace environnemental », les positions des niches modélisées dans cet espace nous informent quant à leurs affinités ou leurs éloignements en terme d'écologie. Nous qualifions alors la position de la niche de « position écologique ». La taille des ellipses est relative à la largeur écologique des niches. Alliée à la position écologique, elle nous renseigne sur le degré relatif de spécialisations des niches ; par exemple : une ellipse réduite en périphérie équivaut à une niche écologique spécialisée. La proportion de chevauchement des ellipses est quant à elle révélatrice du taux de superposition des niches. De cette façon, si les populations modélisées occupent le même territoire au même instant ce chevauchement indique la possibilité d'interactions écologiques, les situe dans l'espace écologique et délivre une notion de leur amplitude.

Enfin, la caractérisation des positions écologiques et des tailles des ellipses d'inerties correspondant à diverses niches d'une même espèce ou d'une même culture modélisée à différentes époques nous permet d'identifier des phénomènes évolutifs comme le décalage, la contraction ou l'expansion de niche.

Synthèse du chapitre : Analyse des niches écologiques

L'observation des projections géographiques fournit aux modélisateurs des informations quant à la distribution potentielle des cultures ou des espèces modélisées. Ainsi, ils peuvent par exemple : discuter des distributions théoriques précédemment établies ou mettre en évidence des similitudes avec l'environnement.

Les analyses multivariées apparaissent comme étant des outils performants pour analyser les niches écologiques. Les ACP individuelles d'une prédiction fournissent des informations sur la composition des niches et quantifient les variables environnementales dans l'établissement du modèle. Les ACP communes permettent, quant à elles, de comparer et de positionner relativement les niches entre elles. Cette analyse menée sur plusieurs prédictions se substitue aux analyses indépendantes de *breadth* et d'*overlap*. Elle permet également d'apprécier l'évolution d'une niche écologique ou écoculturelle en mettant en évidence des expansions, des contractions ou des décalages de niches.

TROISIÈME PARTIE

RECONSTRUCTIONS DE NICHES ÉCOCULTURELLES, TROIS CAS D'ÉTUDES

Impact de l'environnement sur la diversité linguistique de Nouvelle-Guinée

Préambule

Ce chapitre présente une étude menée lors du projet Ethnopaure¹ dirigé par Marian Vanharen. Ce projet consiste à établir des relations entre cultures et biogéographie humaine avec pour objectif l'identification de marqueurs culturels témoignant de ces connexions. Cette approche se veut ethnoarchéologique. Ainsi, nous calibrons notre méthode en analysant les relations entre des traits culturels actuels bien identifiés, leur localisation et leur environnement. De manière à pouvoir étendre par la suite, au travers l'analyse de vestiges archéologiques, cette méthode en différents lieux et à différentes périodes de l'humanité. Dans cette optique nous avons reconstruit les niches écolinguistiques de Nouvelle-Guinée afin d'apprécier la corrélation entre l'environnement et la diversité linguistique de cette région.

Sommaire

8.1	Relations langue/ethnicité/environnement	132
8.2	Contexte environnemental de Nouvelle-Guinée	135
8.3	Histoire du peuplement de la Nouvelle-Guinée, environnement et diversité humaine	139
8.4	Contexte linguistique de Nouvelle-Guinée	146
8.5	Modélisation des niches écolinguistiques de Nouvelle-Guinée	148
8.6	Perspectives	183

1. LabEx Sciences archéologiques de Bordeaux, programme financé par l'ANR - n° ANR-10-LABX-52.

8.1 Relations langue/ethnicité/environnement

La diversité linguistique peut s'expliquer par l'imbrication de phénomènes relevant de dimensions : géographique, temporelle, sociologique et environnementale.

8.1.1 Diversité linguistique/ethnicité

En toute région du monde, on constate des variations linguistiques qui se manifestent par l'apparition de mutations spécifiques. Ces nouvelles langues « naissantes », issues de la somme de ces mutations locales, sont dénommées, selon leur ampleur géographique, géolecte, topolecte, ou régiolecte [Bulot, 2004].

De même, il s'opère un cumul de mutations au fil des générations conduisant à différents chronolectes. À plus grande échelle, ces deux phénomènes de divergence linguistique sont induits par l'isolement physique des populations sur le long terme [Cavalli-Sforza *et al.*, 1988]. Ainsi, il a été longtemps accordé un rôle majeur à la topographie ou à l'insularité dans la divergence des langues (par exemple, [Mabogunje et Richards, 1979 cités par Nettle, 1996], explique que le relief peut constituer un refuge pour certaines populations). De même, réuni autour, et par une même activité, un groupe peut tendre à l'emploi d'un lexique spécialisé pouvant aboutir à la création d'un sociolecte souvent incompréhensible par d'autres groupes menant d'autres activités. Élargi à des sociétés fondées sur des modes de subsistances particuliers, ce mécanisme psychosocial peut conduire à de grandes différences culturelles poussant ces sociétés à se démarquer volontairement les unes des autres. Ces distinctions se faisant sur des critères arbitraux, le phénomène est appelé pseudo-spéciation [Erikson, 1985]. Il est le témoin d'un sentiment d'ethnicité fort. Une langue peut donc être considérée comme un véritable marqueur d'ethnicité au même titre qu'un élément de la culture matérielle, une religion ou un symbole.

Par ailleurs, le phénomène opposé (perte de diversité) peut-être la conséquence d'événements historiques (migrations, variations des taux de mortalité/natalité, etc.) d'origines sociales ou induits par des changements environnementaux. Ainsi, nous constatons que l'étude de la diversité linguistique est complexe et notons la place importante tenue par l'environnement dans le maintien ou l'abandon d'une langue.

8.1.2 Diversité linguistique/environnement

Le langage offre aux humains la possibilité de mieux exploiter leur environnement car il permet la coopération entre individus. Cependant, dans un contexte compétitif, la communication ciblée ou restreinte à un nombre d'individus peut conférer un avantage. De cette manière, la mise en place d'une diversité linguistique dans une région, complexifiant l'intercompréhension entre des groupes d'individus, peut être le résultat d'un processus évolutif en relation avec l'accès aux ressources et donc avec l'environnement.

La diversité culturelle semble être également liée à l'environnement biotique comme le montrent, par exemple, les études de [Cashdan \[2001\]](#), [Collard et Foley \[2002\]](#) ou [Fincher et Thornhill \[2008\]](#) qui indiquent que la densité ethnique augmente avec la densité de maladies infectieuses. Certains auteurs parlent de coévolution entre les diversités culturelles, linguistiques et biologiques [[Smith, 2001](#)]. Cette théorie semble être vérifiée par des observations de variations de densités et d'aires de répartitions suivant un gradient latitudinal commun à ces trois diversités.

8.1.2.1 Le gradient de diversité latitudinal

[Stevens \[1989\]](#) met en évidence un gradient latitudinal de la richesse spécifique, dans les milieux terrestres comme dans les milieux marins. En effet, il remarque que le nombre d'espèces varie en fonction de la distance à l'équateur. De plus, l'auteur souligne que la règle de Rapoport [[Rapoport, 1982](#)], qui stipule que l'aire de répartition d'une espèce diminue avec la latitude, se vérifie également. [Wiens et Donoghue \[2004\]](#) énoncent plusieurs raisons pour expliquer la haute diversité dans les basses latitudes. Premièrement, bon nombre de groupes sont originaires de ces régions, donc leurs histoires évolutives sont plus anciennes que dans les autres régions et ont pu conduire à des divergences plus nombreuses. Deuxièmement, ces auteurs font part de l'hypothèse du « conservatisme tropical » pour expliquer la haute diversité dans les basses latitudes [[Wiens, 2004](#)]. Ils énoncent, par exemple, que les organismes peuvent être, dans certains cas, inadaptés aux températures hivernales. Enfin, ils soulignent que la surface de la zone tropicale s'est considérablement réduite depuis 30 ma², ceci expliquant l'origine tropicale de nombreux *phyla* et le fait que les espèces qui les représentent soient plus nombreuses dans les basses latitudes.

[Mace et Pagel \[1995\]](#) montrent, par une étude sur la diversité linguistique des populations d'Amérique du Nord (pré-contact avec les Européens), que le gradient latitudinal ainsi que la loi de Rapoport peuvent être appliqués aux groupes ethnolinguistiques. Ce constat a été également établi à l'échelle mondiale dans une étude plus récente [[Gavin et Stepp, 2014](#)]. Dans leur étude, [Mace et Pagel \[1995\]](#) constatent également que la diversité linguistique augmente en fonction du nombre d'habitats. Ce dernier point ouvre des perspectives de recherche afin de déterminer quels facteurs environnementaux sont associés aux évolutions linguistiques et culturelles de manière générale.

Des relations identiques à l'échelle de la planète ont été démontrées par l'étude de 3814 cultures humaines [[Collard et Foley, 2002](#)]. De plus cette étude a permis de mettre en évidence des liens entre la diversité linguistique et des facteurs climatiques comme les précipitations et la température. Dans ces travaux, on apprend que les cultures humaines des zones les plus chaudes ou des zones tropicales humides sont moins étalées, sont plus densément regroupées et sont plus diversifiées que dans le reste de la planète.

2. millions d'années

8.1.2.2 Relation entre diversité linguistique et productivité primaire

Quel que soit le mode de subsistance, les ressources disponibles aux Hommes sont dépendantes de la production primaire nette (PPN) qui correspond au bilan de carbone organique produit par un écosystème. Le taux de cette ressource limitante est donc indicateur de la richesse d'un milieu qui détermine les niveaux de compétitions entre les sociétés humaines pour l'accès aux ressources.

Selon une hypothèse, la formation de frontières entre les groupes culturels pourrait être une réponse à la compétition vis-à-vis des ressources [Mace et Holden, 2005]. En ce cas, la diversification des langues serait un des recours pour mettre à bien ce processus. Une autre hypothèse qui ne va pas à l'encontre de la première, stipule que dans un milieu hostile, les Hommes doivent être unifiés et solidaires en élargissant leurs réseaux sociaux et leurs coopérations pour une meilleure exploitation des ressources. Au contraire, de telles contraintes tendent à disparaître lorsque l'environnement est relativement clément [Nettle, 1998]. L'auteur de cette seconde hypothèse spécifie que cette règle est particulièrement vérifiable pour les populations tropicales. Ce résultat est basé sur la corrélation entre densité linguistique par pays et la PPN des écosystèmes dans lesquels ces populations vivent. Dans ses études, Nettle [1996, 1998, 2009] estime cette ressource en calculant la saison de croissance des végétaux (appelée en anglais MGS, pour *mean growing season*)³. Cette droite de régression (figure 8.1) montre que la diversité linguistique croît avec la MGS. Partant du postulat que la MGS est l'inverse du risque écologique, la corrélation permet à Nettle [1996] de vérifier son hypothèse. Nous pouvons critiquer, toutefois, la véracité de l'information qu'apporte cette droite de régression car elle n'est pas établie, selon nous, sur des critères reflétant l'écologie humaine. En effet, pour calculer la densité linguistique, l'auteur se base sur les frontières administratives actuelles, ce qui a peu de sens à nos yeux compte-tenu des bouleversements géopolitiques mondiaux survenus ces derniers siècles.

Selon le même raisonnement, mais agrémenté de données démographiques et génétiques, Coddington et Jones [2013] établissent que la géographie linguistique de la Californie (avant l'arrivée des Européens) aurait été modelée par les migrations de populations pré-historiques. Les auteurs démontrent que les populations s'installent en premier lieu où les conditions environnementales leur sont les plus favorables, puis que d'autres vagues migratrices viennent occuper les habitats plus marginaux. Leur modèle suggère que le sédentarisme accompagné de la capacité à stocker des ressources (conséquence de la néolithisation) contribuent à augmenter la densité de la population initiale et permet le maintien d'une langue dans un territoire. Dans le cas contraire, les épisodes migratoires successifs dans les zones les moins productrices (donc moins peuplées), conduisent au remplacement des populations.

3. nombre moyen de mois convenables à la pousse des végétaux par année

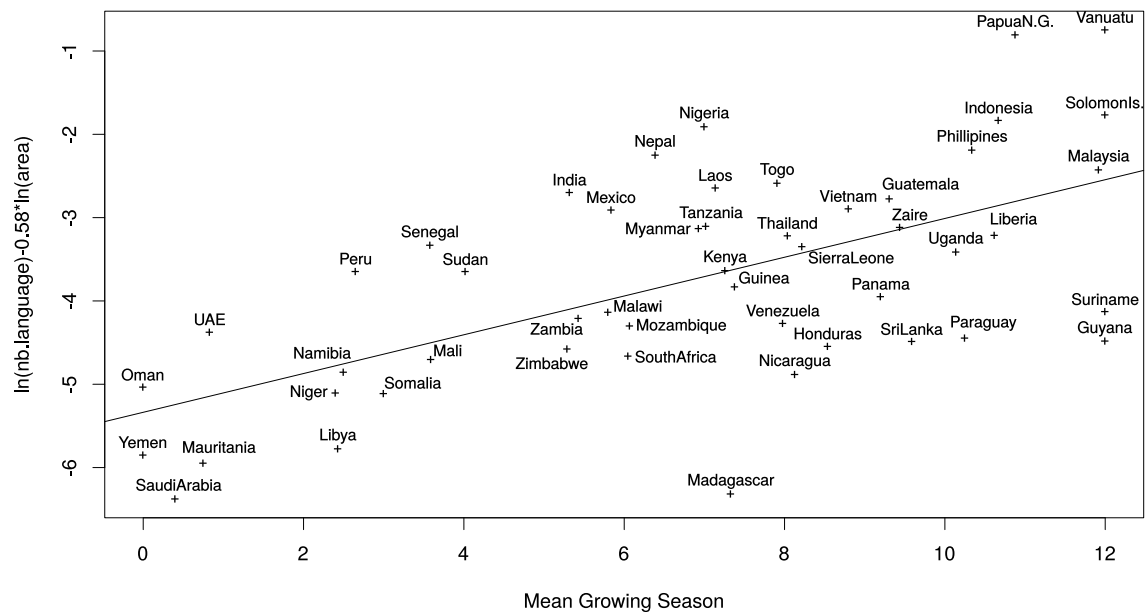


FIGURE 8.1 : Relation entre densités linguistiques et MGS, d'après [Nettle \[1998\]](#).

8.2 Contexte environnemental de Nouvelle-Guinée

Cette section présente l'environnement de la Nouvelle-Guinée, sous ses aspects géographiques et écologiques.

8.2.1 Géographie physique et climat

La Nouvelle-Guinée est la deuxième île la plus grande au monde après le Groenland. Elle est politiquement divisée en deux parties de tailles relativement équivalentes (figure 8.2). La partie occidentale, anciennement dénommée Irian Jaya, est constituée de deux provinces d'Indonésie : la province de Papouasie Occidentale (ou *Papua Barat*) et la province de Papouasie. La partie orientale est un pays indépendant nommé Papouasie-Nouvelle-Guinée. La forme générale de l'île peut être comparée à celle d'un oiseau, la péninsule de Doberai (la région la plus occidentale) est d'ailleurs parfois dénommée *the Bird's Head*⁴. L'île possède une cordillère tout son long qui sépare le Nord du Sud. Cette chaîne montagneuse peut dépasser les 4500 mètres. Dans ce document, la dénomination de Hautes Terres fait référence

4. En français : la tête de l'oiseau.

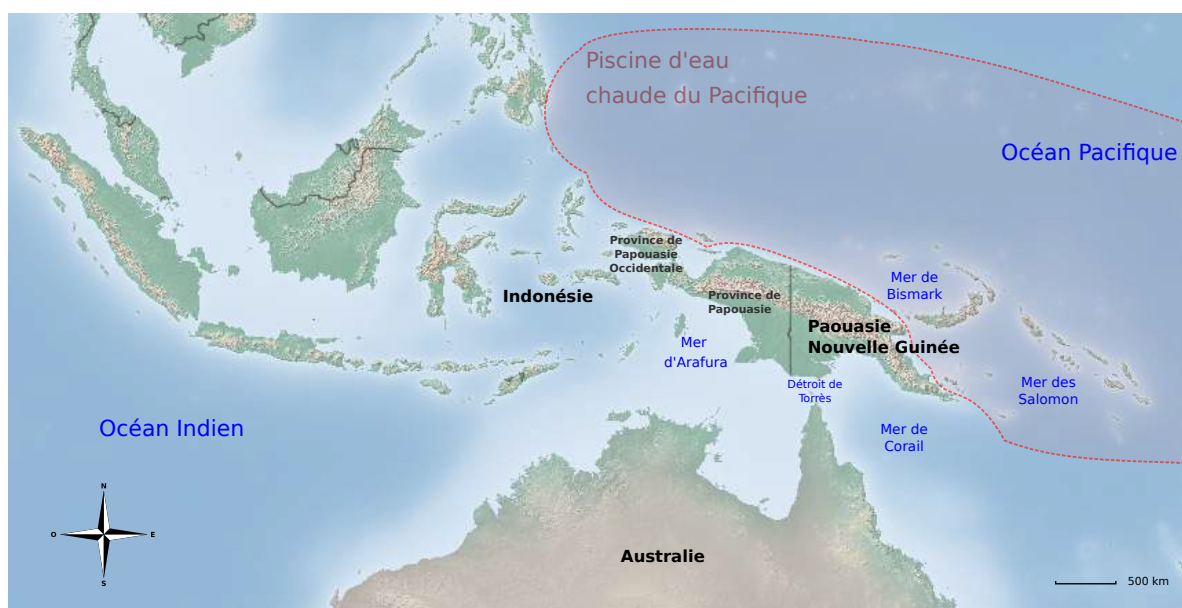


FIGURE 8.2 : Position géographique de l'archipel de Nouvelle-Guinée.

à cette région montagneuse relativement peuplée. Cette île, située en zone équatoriale (entre les latitudes 0° et 11°S), est bordée au Nord par l'Océan Pacifique ainsi que la mer de Bismarck, à l'Est par la mer des Salomon et au Sud par la mer chaude et peu profonde d'Arafura et la mer de Corail (voir figure 8.2).

Proche de l'équateur, au Nord du continent austral, l'île se situe dans la « zone des calmes tropicaux », ce qui lui confère un climat relativement stable tout au long de l'année. Néanmoins, des régions au Sud et en Papouasie-Nouvelle-Guinée comportent une saison sèche de Mai à Octobre. La Nouvelle-Guinée est le prolongement septentrional de la plaque Australe qui forme une barrière entre le Pacifique Ouest et l'Est de l'Océan Indien. Il en résulte la formation de la Piscine d'eau chaude du Pacifique Ouest [Prentice et Hope, 2007] (voir figure 8.2), qui provoque l'oscillation el Niño / la Niña, créant une mousson dans ces régions. De manière générale, les précipitations sont plus importantes au Nord qu'au Sud (par exemple 1010 mm annuel à Port Moresby, contre 3696 mm à Kawaliap dans l'île de Manus⁵). D'après la classification de Köppen-Geiger [Kottek *et al.*, 2006], on trouve en Nouvelle-Guinée principalement des climats du groupe A (climats tropicaux) :

Af Climat tropical humide, caractérisé par des précipitations abondantes toute l'année, aucune saison n'est particulièrement marquée. Ce climat est le plus répandu en Nouvelle-Guinée (exemple figure 8.3(c))

5. données relevées sur www.climate-data.org

Am Climat tropical avec mousson, on le trouve au Sud et à l'Est de la Nouvelle-Guinée (exemple figure 8.3(b))

Aw Climat tropical avec une saison sèche, localisé au Sud et au Sud-Est de la Nouvelle-Guinée (exemple figure 8.3(d))

Cfb Climat tempéré océanique chaud sans saison sèche, peu répandu (exemple figure 8.3(e))

Sous l'influence des différentes mers et océans l'entourant ainsi que de son relief, l'île présente des milieux très diversifiés.

8.2.2 Diversité environnementale de Nouvelle-Guinée

La WWF⁶ définit une éco-région terrestre comme l'ensemble des lieux d'une même zone possédant des caractéristiques géographiques et des assemblages d'espèces communs. Grâce à de multiples collaborations, l'organisation a dressé une base de données cartographique de ces zones et des biomes auxquels elles appartiennent [Olson *et al.*, 2001]. À la suite de l'interrogation de cette base, nous avons relevé, pour la région qui correspond à la zone d'étude, 520 écosystèmes terrestres pouvant être regroupés en 26 éco-régions (figure 8.4). Ces éco-régions sont des subdivisions de 4 biomes (figure 8.5).

Tropical & Subtropical Grasslands, Savannas & Shrublands (prairies, savanes, brousses tropicales) : les précipitations faibles dans ces régions produisent une végétation principalement herbacée ou arbustive.

Tropical & Subtropical Moist Broadleaf Forests (forêts déciduales humides tropicales et subtropicales) : le climat chaud et très humide permet un couvert forestier dense constitué d'espèces d'arbres de grandes tailles. La richesse spécifique de la faune et de la flore y est plus élevée que dans les autres biomes. On y trouve des forêts primaires aussi appelées forêts équatoriales ou forêts pluviales.

Mangrove : situé dans les zones intertidales⁷, elles sont constituées majoritairement de palétuviers.

Montane Grasslands & Shrublands (prairies et brousses d'altitude) : la végétation est entièrement constituée d'herbes et d'arbustes, ce biome est aussi dénommé pelouses alpines. Le climat y régnant est froid et humide, ces prairies constituent des isolats qui hébergent des espèces endémiques.

La Nouvelle-Guinée étant située en zone équatoriale et subéquatoriale, les durées des jours et des nuits sont approximativement identiques tout au long de l'année. Le biome le plus présent est la forêt équatoriale (figure 8.5), le climat y régnant est chaud et très humide (pluies continues), avec peu d'amplitude thermique journalière et saisonnière.

6. World Wildlife Fund

7. zones de balancement des marées

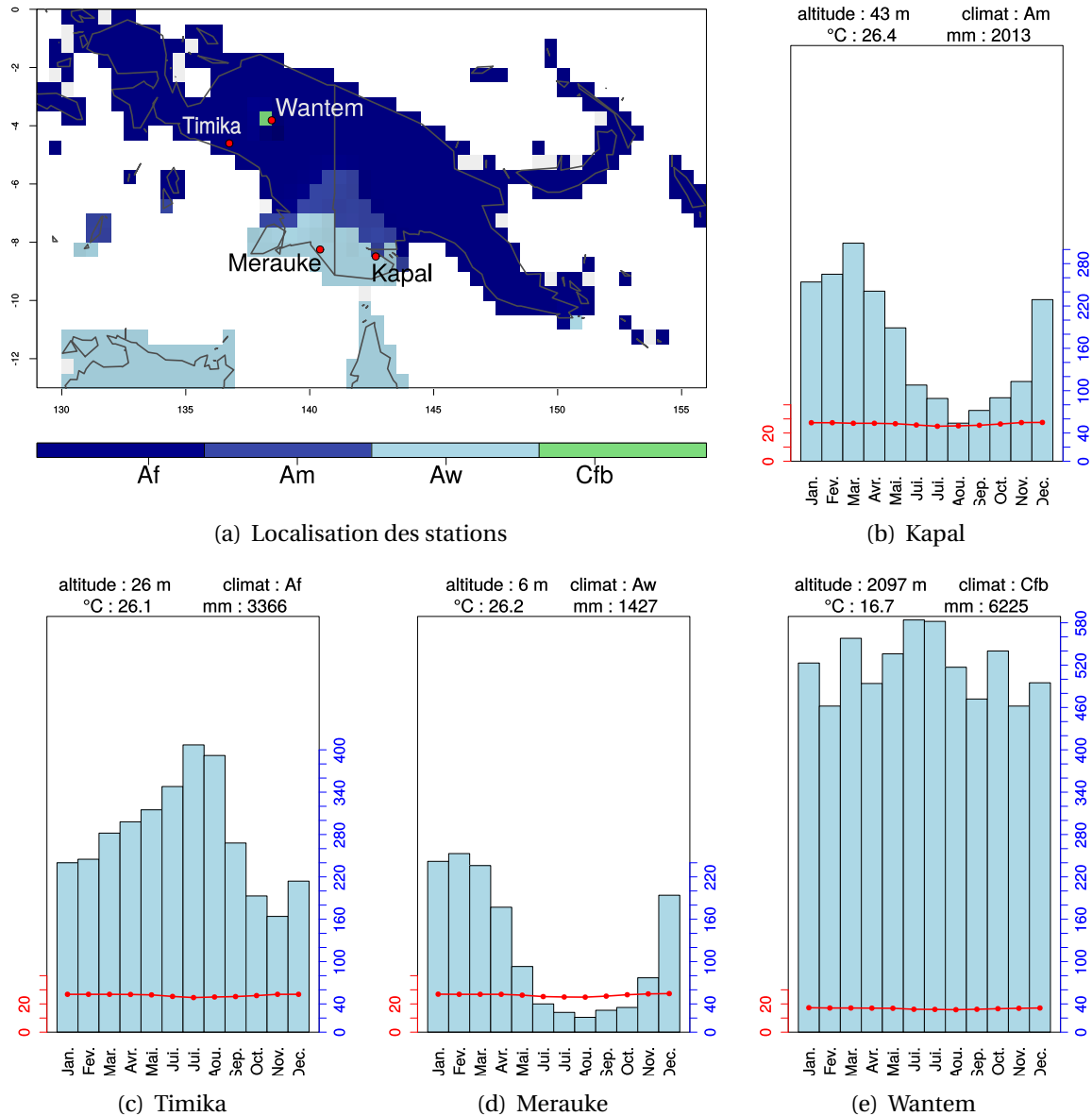


FIGURE 8.3 : Climats de Nouvelle-Guinée et diagrammes ombrothermiques des villes de Kapal, Timika, Merauke et Wantem. Pour chaque station sont indiqués : l'altitude, le climat (selon la classification de Köppen-Geiger), la température moyenne annuelle (°C) et les précipitations annuelles (mm). Les courbes représentent les températures mensuelles (axe de gauche), les histogrammes : les précipitations mensuelles (axe de droite). Les diagrammes ont été réalisés à partir des données climatiques disponibles sur la base de données en ligne : www.climate-data.org.

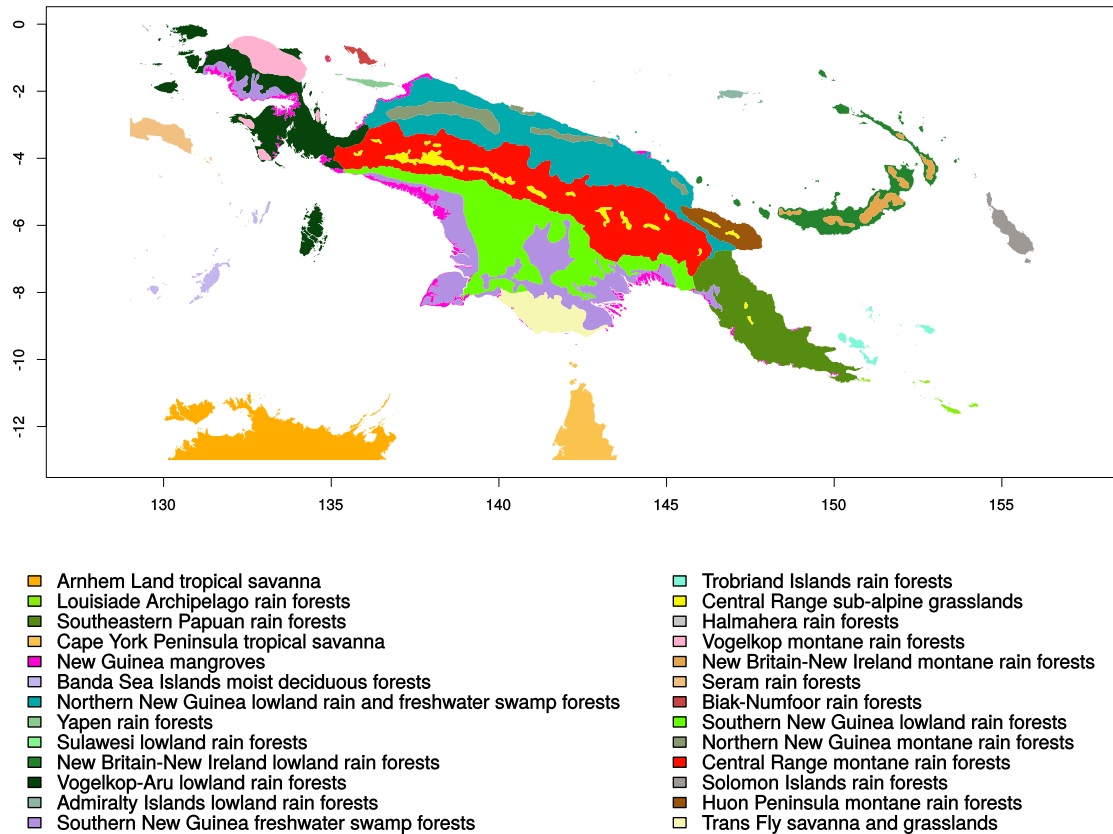


FIGURE 8.4 : Carte des éco-régions. Cette carte a été obtenue à partir du fichier de la WWF qui recense les écosystèmes terrestres [Olson *et al.*, 2001].

Dans cette étude nous allons confronter diversité environnementale et diversité linguistique. Ainsi nous vérifierons si une étude écolinguistique de la Nouvelle-Guinée peut valider l'hypothèse que nous énoncions dans le chapitre 1, à savoir que la complexité de la géographie linguistique (ou culturelle au sens large) trouve ses racines dans l'histoire des populations et leur adaptations à l'environnement.

8.3 Histoire du peuplement de la Nouvelle-Guinée, environnement et diversité humaine

La linguistique, la génétique et l'archéologie s'accordent, toutes trois, sur l'hypothèse d'un peuplement de la Nouvelle-Guinée par séries de vagues [Bellwood, 1991; Capelli *et al.*, 2001; Cavalli-Sforza *et al.*, 1994]. La première série peuple l'Australie il y a plus de

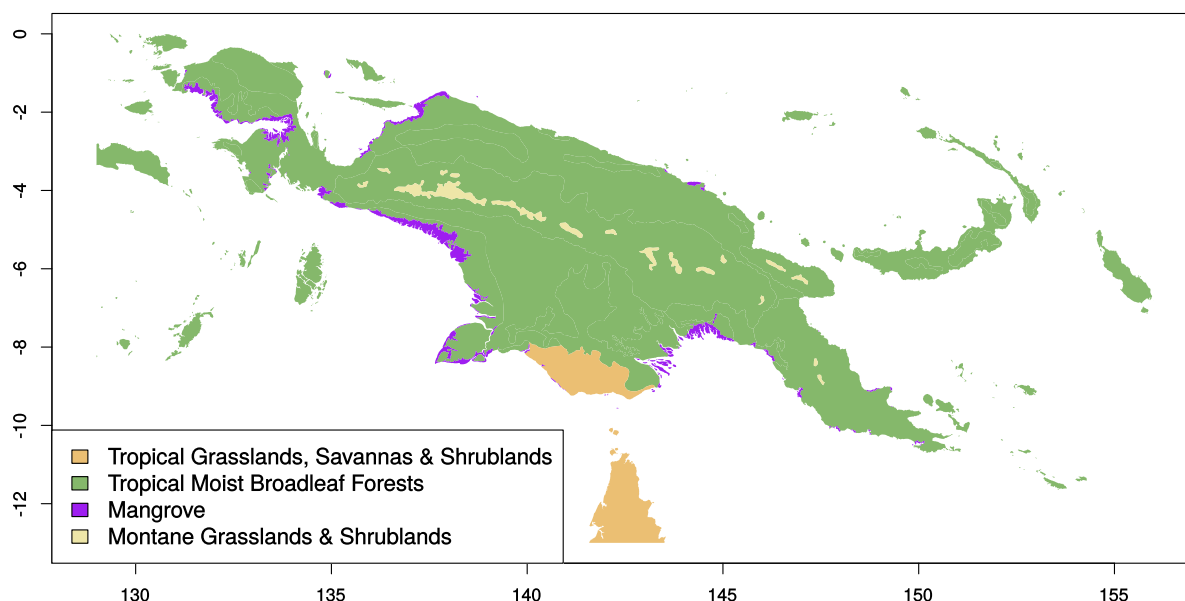


FIGURE 8.5 : Carte des biomes de Nouvelle-Guinée. Cette carte a été obtenue à partir du fichier de la WWF qui recense les biomes terrestres [Olson *et al.*, 2001].

60-50 ka ans [Roberts *et al.*, 1990] puis la Mélanésie il y a environ 50-40 ka [Allen, 1996; O'Connell et Allen, 2004; Summerhayes *et al.*, 2014]. Les migrations humaines se trouvent alors facilitées dans certaines régions par le bas niveau des mers. Cette première série de vagues en Nouvelle-Guinée constitue le peuplement papou. La deuxième série, dont les dates et les lieux d'émergences varient selon les auteurs, est plus tardive. Il s'agit du peuplement austronésien. Ces derniers possèdent une technologie de navigation plus avancée que leurs prédécesseurs leur permettant de s'étendre sur deux océans. Ils atteignent Madagascar dans l'Océan Indien et l'Île de Pâques dans le Pacifique [Bellwood, 1991; Pawley, 2002].

8.3.1 Le peuplement papou

La Nouvelle-Guinée, du fait de son histoire géologique, n'a pas toujours été isolée. En effet, les fluctuations marines ont pu jouer un rôle dans le peuplement de cette île ainsi que dans la diversification culturelle et biologique de ses populations humaines.

8.3.1.1 La variation du niveau marin

Les premiers peuplements de la Nouvelle-Guinée se sont probablement effectués suivant deux routes depuis les anciennes terres émergées du Sunda qui rattachaient les îles de

Java, Bali, Borneo et Sumatra au continent asiatique via la presqu'île Malaise [Schiefenhövel, 2014] (voir figure 8.6). L'une de ces routes traversait Java et Timor avant de rejoindre le

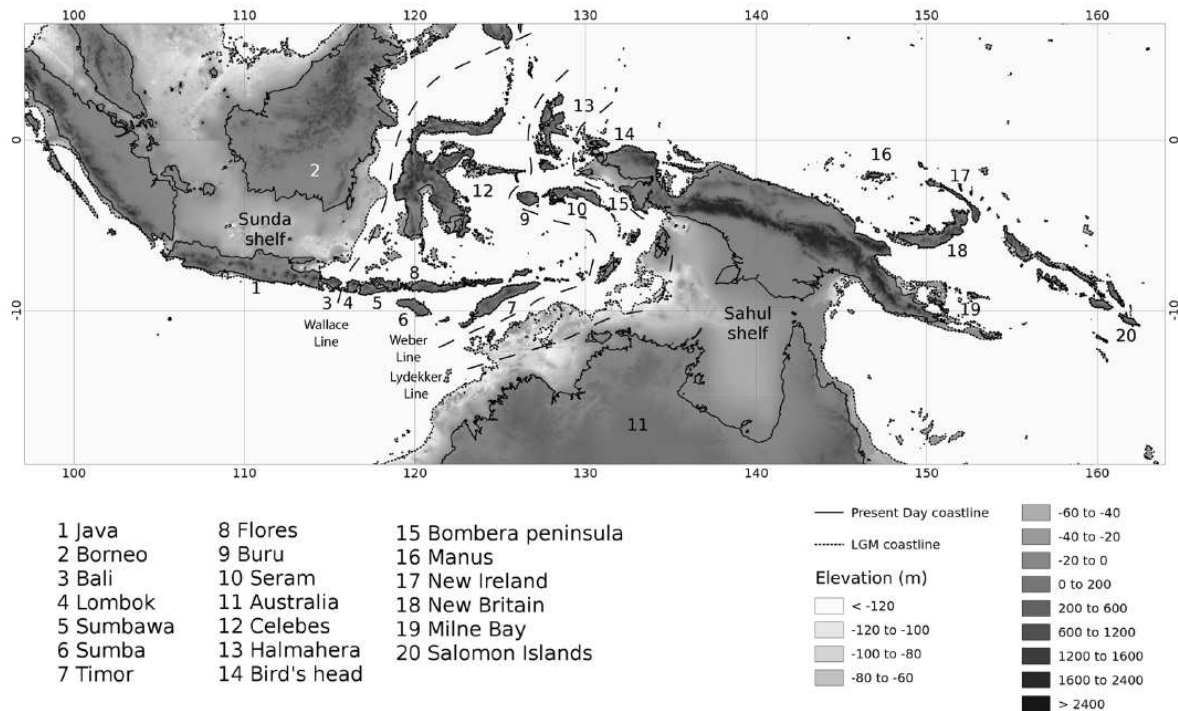


FIGURE 8.6 : Océanie au DMG, d'après [Schiefenhövel, 2014]

Sahul. L'autre route, plus septentrionale, cheminait par Bornéo et Celebes (*Sulawesi*, où ont été découvertes récemment des peintures rupestres datant d'environ 40 ka [Aubert *et al.*, 2014]) avant de gagner la Province de Papouasie Occidentale par la Péninsule de Doberai (*Bird's Head*). Bien que le niveau marin ait été plus bas par le passé (jusqu'à 120 mètres au dessous du niveau actuel durant le dernier maximum glaciaire, figure 8.6), les déplacements entre certaines îles se trouvant sur les deux routes ont de tout temps nécessité la possession d'embarcations comme des pirogues ou des radeaux. La circulation de la Nouvelle-Guinée à l'Australie, quant à elle, pouvait s'effectuer par voie terrestre, car le Sahul-Arafura est longtemps resté une terre émergée. Cet isthme a certainement favorisé des migrations humaines accompagnées d'échanges culturels jusqu'à son immersion complète il y a environ 8 ± 0.1 ka [Chappell, 2005]. Ainsi la naissance de la mer d'Arafura divise le Sahul en deux parties distinctes, au Sud l'Australie et au Nord la Nouvelle-Guinée, se différenciant aussi bien géographiquement qu'écologiquement (voir figure 8.4).

Longtemps connectées, les terres de l'ancien Sahul partagent aujourd'hui encore une partie de leur faune et de leur flore, cependant de nombreuses espèces sont endémiques de

certaines îles. Le détroit de Torrès (figure 8.2) dresse une véritable barrière qui engendre des différenciations biologiques [Kayser *et al.*, 2001] et culturelles [Hays, 1994] entre l'Australie et la Nouvelle-Guinée.

8.3.1.2 L'ethnoécologie papoue

Les premiers Papous. À la fin du Pléistocène, les Papous qui colonisèrent pour la première fois les terres de Nouvelle-Guinée étaient des chasseurs cueilleurs, à ce jour nous avons peu de connaissances relatives à ces groupes humains car peu de sites archéologiques ont été fouillés [Dennell et Porr, 2014; Schiefenhövel, 2014; Summerhayes *et al.*, 2014]. Nous pouvons citer néanmoins les sites archéologiques de Bobongara (en péninsule Huon daté à plus de 44 ka [Fairbairn *et al.*, 2006]), de Kosipe (situé à l'Ouest du Mont Albert Edward, ancien d'au moins 35 ka [White *et al.*, 1970]), de Nombe (25 ka, dans la province de Simbu [Mountain, 1993]), de FABM (en Nouvelle Bretagne, entre 27 et 40 ka [O'Connell et Allen, 2004; Torrence *et al.*, 2004]) et de Buang Merabak (nouvelle Irlande, daté d'environ 40 ka [Fairbairn *et al.*, 2006; O'Connell et Allen, 2004]). Ces deux derniers sites de l'archipel Bismarck pourraient suggérer une troisième voie maritime le long de la côte Nord de la Nouvelle-Guinée qui aurait débuté avant l'Holocène.

Les autres preuves de leur présence sur l'île à cette époque sont fournies de façon indirecte par la génétique et la linguistique essentiellement. Des études palynologiques [Golson, 1991] ou anthracologiques [Haberle, 2003; Haberle *et al.*, 2001] nous renseignent sur les modes de subsistance de ces groupes humains. En effet, les populations vont, selon leurs localités, se spécialiser différemment. Les populations côtières vont se tourner vers la pêche, celles des zones marécageuses vers la chasse et dans les Hautes Terres (entre 700 et 3000 m) les populations vont, à l'Holocène moyen, domestiquer des espèces végétales et entamer un processus de néolithisation.

Un foyer de néolithisation. Il y a environ 8 ka, l'horticulture et l'arboriculture apparaissent en Nouvelle-Guinée [Golson, 1976]. Il s'agit d'une découverte autonome, accompagnée de technologies comme la pierre polie, de la domestication d'espèces végétales et du sédentarisme; on peut alors parler de foyer de néolithisation comparable à celui du Proche-Orient il y a 11.5 ka ou ceux, plus récents, de Mésoamérique, d'Afrique du Nord-Est ou de Chine (ces néolithisations ne sont pas toutes complètes, mais attestent au moins de la domestication d'espèces végétales [Bower, 1991; Crawford, 2006; Crawford *et al.*, 2005; Piperno, 2011; Purugganan et Fuller, 2009; Smith, 2011]). Pour travailler la terre, les Papous utilisent des bâtons fouisseurs, des bêches en bois et des herminettes en pierre polie [Di Piazza, 1996; Pétrequin et Pétrequin, 1993]. Ils ont domestiqué la canne à sucre (*Saccharum officinarum*), plusieurs espèces de taro (*Colocasia esculenta*, *Cyrtosprema chamissonis*, *Alocasia macrorrhiza*), l'ignam (*Dioscorea bulbifera*) ainsi que de nombreuses autres espèces végétales comme des arbres fruitiers [Schiefenhövel, 2014, 2013]. Afin de compléter l'apport protéique journalier, il est pratiqué dans la population actuelle des Eipos,

membre du groupe linguistique Mek (situé dans les Hautes Terres), la consommation de petits animaux chassés au collet et d'insectes [Schiefenhovel et Blum, 2007].

Un mode de vie traditionnel. Une grande partie des populations papoues vivent en montagne pour éviter les zones impaludées. Cette particularité géographique et la découverte relativement récente de la région par les Européens (xvi^e siècle), font que nombre de ces populations ont été préservées de contacts directs avec le mode de vie occidental jusqu'à il y a encore peu de temps.

Si le contact avec les occidentaux est récent et a transformé depuis peu leurs sociétés, les cultures papoues, notamment celles occupant la périphérie de l'île principale, sont en contact depuis plusieurs millénaires avec les Austronésiens. Cette proximité est attestée par l'introduction de nombreux éléments dans leurs cultures comme l'élevage porcin [Oppenheimer et Richards, 2001].

8.3.2 Le peuplement Austronésien

La diffusion des Austronésiens à travers le monde est parmi les plus récentes des expansions majeures préhistoriques [Kayser *et al.*, 2006]. Ayant eu lieu il y a quelques millénaires et ayant été peu altérées par les migrations modernes, les structures génétiques et linguistiques de ces populations, confortées par les études archéologiques (principalement axées sur la comparaison du mobilier et des techniques agricoles comme les espèces domestiquées) [Capelli *et al.*, 2001], permettent de retracer l'histoire de ce peuplement avec un degré de fiabilité relativement élevé. Ces études semblent indiquer que ces populations ont atteint la Nouvelle-Guinée et les îles environnantes il y a environ 4000 ans et seraient originaires de Taïwan [Lynch *et al.*, 2002; Melton *et al.*, 1995] ou de l'Asie du Sud-Est [Bellwood, 1978]. Depuis les années 2000, plusieurs théories concernant les dates, la vitesse d'expansion et les origines des Austronésiens s'affrontent.

8.3.2.1 Les différentes théories du peuplement austronésien

La théorie de l'express-train [Bellwood, 1985; Kirch, 1997], popularisée par Diamond [1988], est l'hypothèse de peuplement austronésien la plus connue. Essentiellement fondée sur les vestiges archéologiques, comme les tessons de céramique lapitas⁸, de l'obsidienne de Talasea, ainsi que des preuves d'activités agricoles et d'élevage. Cette hypothèse suggère un peuplement rapide (moins d'un millénaire) depuis Taïwan vers les îles éloignées du Pacifique oriental il y a 6000 ou 3000 ans. L'hypothèse stipule que les migrants ne se seraient pas mélangés aux populations papoues alors autochtones. De nos jours, des preuves archéologiques [Noury et Galipaud, 2011], linguistiques

8. La culture Lapita est caractérisée par une céramique pointillée à décors incisés et reliefs appliqués. Cette culture est retrouvée en plusieurs lieux de Nouvelle-Guinée et de Polynésie, les premiers vestiges lapitas ont été découverts dans l'archipel Bismark il y a 3500 ans [Green, 2000].

[Holzknecht, 1989; Schiefenhövel, 2014] et biologiques [Kayser *et al.*, 2006, 2000] vont à l'encontre de cette théorie de peuplement rapide et sans escale.

La théorie du *slow-boat* [Kayser *et al.*, 2000], évoque une migration austronésienne lente originaire du sud est de la Chine et de Taïwan à partir de 6000 BP. Ce modèle s'appuie sur des études d'ADN mitochondrial et du chromosome Y qui ont conduit les chercheurs à observer que des haplogroupes d'ADN mitochondrial pouvaient être retrouvés tout au long de la Mélanésie alors que les fréquences haplotypiques du chromosome Y suggéraient, dans la plupart des cas, une spécificité locale [Kayser *et al.*, 2006, 2003, 2008]. Archéologiquement, le modèle est conforté par de nombreuses similitudes culturelles retrouvées sur toute la région. Lors de cette expansion lente, les austronésiens auraient véhiculé (mais aussi laisser derrière eux) des gènes papous ainsi que des technologies et des espèces indigènes [Oppenheimer et Richards, 2001].

La théorie des Nusantao : Solheim [1984] n'explique pas la présence de traits culturels communs aux différentes entités culturelles austronésiennes par les migrations. Il conçoit plutôt des populations tournées vers la mer appartenant à un vaste réseau d'échanges et ayant eu la nécessité de développer une intercompréhension pour favoriser le commerce. L'auteur parle alors de Nusantao pour désigner ce groupe de populations. Sa théorie est principalement fondée sur l'absence d'un gradient de traits culturels parmi les vestiges archéologiques. Il estime d'ailleurs que plusieurs familles, non nécessairement austronésiennes, sont à l'origine du Nusantao. Selon l'auteur, l'origine de ce réseau se situerait il y a 7000 ans au Nord de Luçon (Philippines), au Sud de Taïwan et en Chine du sud.

La théorie du triple I (*Intrusion, Innovation, Integration*) [Allen, 2000]. Ce modèle émet l'hypothèse que l'« intrusion » de populations provenant d'Asie du Sud-Est permet, par de nombreux échanges et contacts, l'« intégration » de techniques et d'éléments culturels induisant l'« innovation » caractérisée par un long processus de changement de mode de subsistance [Green, 2000; Noury et Galipaud, 2011].

La théorie *pulse-pause* [Gray *et al.*, 2009]. Ce modèle explique que le peuplement austronésien s'est déroulé par phases stationnaires longues entrecoupées d'expansions rapides. La première « pause » concernerait la région de Taïwan entre 5000 et 6000 BP et serait suivie d'une extension rapide sur toute la Malaisie, la Micronésie et la partie occidentale de la Polynésie. Ce modèle énonce aussi la possibilité d'une deuxième extension provenant de l'Est de l'Indonésie vers 3000 BP. Cette hypothèse suggère que lors des « pauses », les populations échangent des gènes ainsi que des éléments culturels, rejoignant ainsi l'hypothèse du triple I.

8.3.2.2 Le mode de vie austronésien

Les austronésiens sont une société de marins, et sont, comme les Papous des horticulteurs. Cependant, ils ne cultivent pas les mêmes espèces, car ils ont hérité de celles

domestiquées par les populations du Sud-Est asiatiques. Leur maîtrise de l'art de la céramique leur permet le stockage et le transport de ces végétaux. Ils pratiquent également l'élevage de volailles, de porcs et probablement de chiens. Principalement localisés dans les îles périphériques de la Nouvelle-Guinée ou en régions côtières de l'île principale, les Austronésiens sont fortement exposés au paludisme.

8.3.3 Impact de l'environnement biotique sur les génotypes humains et le peuplement

Les écosystèmes de Nouvelle-Guinée et d'Australie hébergent une grande diversité d'espèces végétales et animales dont la plupart sont caractéristiques de l'écozone australasienne⁹ et d'autres spécifiques à la région comme les marsupiaux. Mais ce climat favorable au développement de la flore et de la faune l'est aussi aux agents pathogènes. La forte humidité ambiante tout au long de l'année est propice aux moustiques vecteurs de maladies telles la dengue, le paludisme, la framboisie, l'*Ulcus tropicum* et plusieurs formes de helminthiase [Braun, 1996; Jüptner, 1983; Schiefenhövel, 1982].

Parmi les quatre espèces de parasites responsables du paludisme présentes en Nouvelle-Guinée, deux d'entre elles : *Plasmodium vivax* et *Plasmodium falciparum*, sont majoritaires. L'espèce humaine, face à la forte prévalence du paludisme, sélectionne des gènes, bien que responsables de maladies, comme le gène du favisme qui suscite aux porteurs un déficit en glucose-6-phosphate déshydrogénase mais dont l'allèle à l'état hétérozygote, prévient des symptômes du *Plasmodium falciparum* [Kirkman *et al.*, 1968; Ruwende et Hill, 1998]. La fréquence de cet allèle peut atteindre 53% dans certaines populations de Nouvelle-Guinée [Müller *et al.*, 2003]. De la même manière la mutation affectant le troisième exon du gène de la Glycophorine, C qui confère au porteur le groupe sanguin Gerbich-négatif [Booth et McLoughlin, 1972], est favorisée car elle réduit les effets de *Plasmodium vivax* et *Plasmodium falciparum*. Les populations d'altitude vivant au-dessus de 1300-1600 m ne sont pas exposées au paludisme car elles se trouvent hors de l'habitat de l'anophèle, moustique vecteur du paludisme. Par conséquent, la fréquence allélique de la mutation de Gerbich y est relativement faible alors qu'elle avoisine les 10% dans les terres de basse altitude [Müller *et al.*, 2003].

Les migrations humaines peuvent expliquer en partie la diversité humaine [Cavalli-Sforza *et al.*, 1994; Kayser *et al.*, 2003; Stoneking *et al.*, 1990]. Cependant, les cas du favisme et du groupe sanguin de Gerbich, qui ne sont que deux exemples parmi d'autres, nous amènent à constater que l'histoire évolutive et l'environnement modèlent également les structures génétiques des populations. Ainsi, nous voyons grâce ces exemples que via la sélection naturelle, les individus porteurs de certaines mutations génétiques sont alors favorisés et que les populations auxquelles ils appartiennent ont la possibilité de se développer dans des

9. regroupement d'éco-régions d'Australie et d'Asie du Sud-Est

lieux impaludés. Avec elles, ces populations véhiculent d'autres gènes, mais aussi, d'autres systèmes adaptatifs, d'autres cultures, d'autres langues.

8.4 Contexte linguistique de Nouvelle-Guinée

Plus d'un millier de langues sont parlées dans cette région. On compte près de 800 langues papoues et 300 austronésiennes [Reesink *et al.*, 2009; Wurm, 1983].

8.4.1 Une diversité linguistique très élevée

Les langues évoluent par des altérations, des pertes ou des gains de vocables ou de structures grammaticales, causés par les expériences historiques culturelles ou environnementales [Cavalli-Sforza *et al.*, 1994; Swadling, 1997; Velichkovsky et Rumbaugh, 2013]. Comme pour la biodiversité, la singularité et la richesse linguistique d'un endroit sont dépendantes de son isolement, de l'apport de nouveaux éléments, d'événements historiques ainsi que des interactions entre les populations. En Mélanésie, tous ces ingrédients sont réunis, mais est-ce suffisant pour expliquer cette si grande richesse linguistique ou ce phénomène est-il induit également par un environnement propice à la diversité comme certains auteurs l'annoncent [Coddington et Jones, 2013; Mace et Pagel, 1995; Nettle, 1998] ?

Dans ce chapitre, nous nous efforçons de vérifier ce dernier point grâce à la modélisation de niches de groupes linguistiques appartenant à chacune des deux grandes familles de langues présentes dans cette région : la famille papoue et la famille austronésienne [Wurm *et al.*, 1975].

8.4.2 Phylolinguistique de Nouvelle-Guinée

8.4.2.1 La famille de langues papoues

Les langues papoues trouvent probablement leurs racines dans les langues parlées lors des premiers peuplements du Sahul (voir section 8.3). Cette famille, uniquement présente en Mélanésie et dans les îles de Timor, Alor et Halmahera compte plus de 800 langues [Reesink *et al.*, 2009]. Elle constitue un groupe paraphylétique, probablement la résultante d'une longue et complexe histoire évolutive [Foley, 1986; Golson, 2005; Pawley, 2005; Swadling, 1983]. Les langues papoues sont ainsi divisées en plusieurs groupes : le TNG (*Trans New Guinea* couleurs vertes de la carte figure 8.7) qui comprend plus de 300 langues à lui seul, 22 autres groupes moins conséquents (non-TNG, dans les tons rouges et gris de la figure 8.7), plus quelques isolats [Ross *et al.*, 2005].

Il y a près de 4000 ans, les populations papoues occupant les régions périphériques de l'île sont rentrées en contact avec les Austronésiens (voir section 8.3.2). La rencontre des deux familles de langues vient complexifier d'autant plus la phylogénie linguistique de la région.

8.4.2.2 La famille de langues austronésiennes

Cette famille de langues est la deuxième plus grande au monde¹⁰ avec plus de 1200 langues [Gordon et Grimes, 2005] dont près de 300 sont parlées en Nouvelle-Guinée et les îles environnantes [Reesink *et al.*, 2009]. Les premiers contacts entre Papous et Austronésiens ont eu lieu il y a environ 4000 ans (voir section 8.3.2). Les langues austronésiennes de la région étudiées font exclusivement partie de la sous-famille des Océaniens occidentaux qui est probablement originaire de Taïwan selon Melton *et al.* [1995] et Lynch *et al.* [2002] (en effet de nos jours, dans l'île de Taïwan, sont parlées des langues appartenant à neuf *phyla* sur les dix *phyla* que compte la famille de langue austronésienne [Blust, 1999 cité par Capelli *et al.*, 2001]), alors que selon Bellwood [1978], les langues austronésiennes proviendraient de l'Asie continentale du Sud-Est. Les parties colorées dans les tons bleus de la carte figure 8.7 indiquent des régions où sont actuellement parlées des langues austronésiennes dans notre aire d'étude.

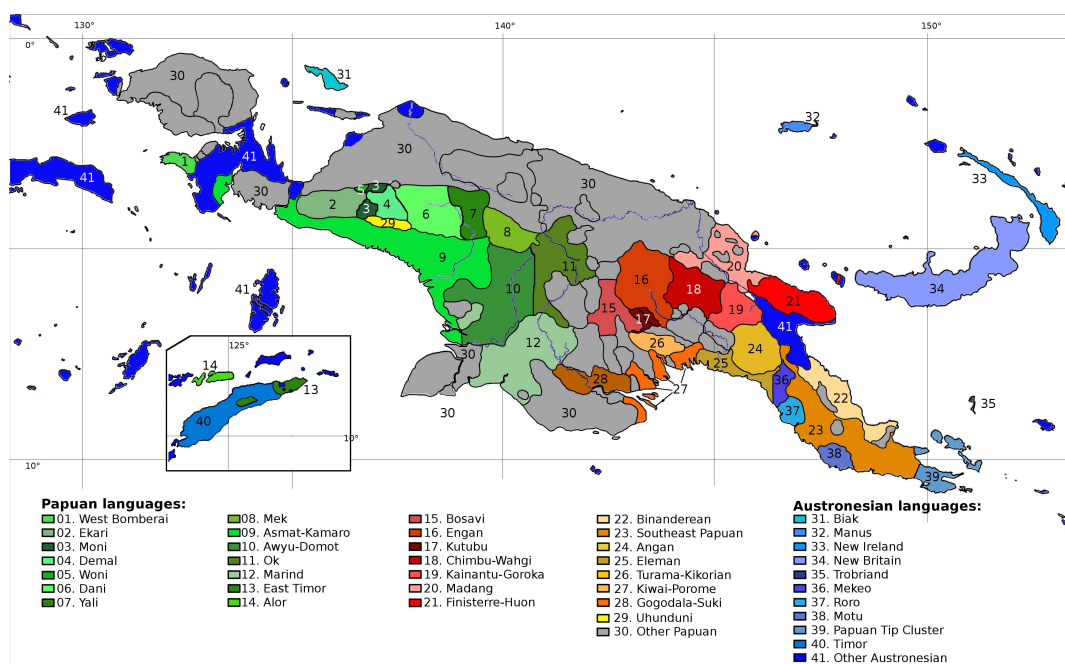


FIGURE 8.7 : Diversité linguistique de Nouvelle-Guinée, d'après Schiefenhövel [2014].

Ces vagues de peuplements austronésiennes ont contribué à l'augmentation de la diversité génétique de la région déjà fortement présente due à la dérive sémantique découlant de l'isolement et d'autres mécanismes évolutifs [Cavalli-Sforza *et al.*, 1994; Swadling, 1997].

Nous venons de voir comment l'histoire du peuplement et l'évolution biologique ont pu contribuer à la grande diversité linguistique de Nouvelle-Guinée. Dans cette thèse nous

10. La famille qui compte le plus de langues est l'Indo-européen

avons cherché à déterminer si la variabilité environnementale a pu également contribuer par la suite à la richesse linguistique de cette région.

Étudier la diversité linguistique en Nouvelle-Guinée. Afin de tester l'hypothèse de la relation langues/environnements dans l'objectif d'extrapoler aux relations cultures/environnements, nous avons conduit une ECNM (section 1.6.1) de différents groupes ethnolinguistiques de Nouvelle-Guinée. Ainsi, nous obtenons des niches écolinguistiques (ELN¹¹) et, à l'instar de l'ECNM, nous avons appelé cette méthode : *Ecolinguistic niche modeling* (ELNM).

8.5 Modélisation des niches écolinguistiques de Nouvelle-Guinée

Une niche écolinguistique (ELN) représente l'étendue des conditions environnementales exploitées par une population parlant la même langue ou des langues très proches. La modélisation de niches écolinguistiques (ELNM) se fait suivant la même procédure méthodologique que l'ECNM. La modélisation a été conduite pour des groupes de langues situés à un même niveau dans la phylogénie [Ross *et al.*, 2005] qui réunissent des langues selon leurs apparentés lexicales et syntaxiques.

8.5.1 Occurrences

Nous avons mené l'ENM selon 26 groupes de langues papoues faisant partie de plusieurs sous-familles, dont la famille TNG (voir section 8.4.2.1) et neuf groupes de langues austronésiennes de la branche océanienne occidentale. Ces groupes sont présentés dans la figure 8.7, où les groupes austronésiens portent les index 31 à 40 (figure 8.7) et les groupes papous les index 1 à 29. Nous n'avons ni modélisé le groupe linguistique *other Papuan* ni le groupe *other austronesian*, car exceptée leur appartenance respective à une même famille, ils constituent chacun un ensemble de langues appartenant à des clades différents, leur regroupement en un unique groupe « autre » n'est donc pas cohérent. De même, nous avons écarté de l'analyse les groupes éloignés géographiquement : *East Timor*, *Timor* et *Alor* car leur éloignement géographique peut fausser la modélisation (voir section 8.5.3). Enfin nous avons dû renoncer à la modélisation du groupe linguistique *Woni* car nous ne disposions pas assez d'occurrences ($n = 7$) pour obtenir des prédictions fiables (voir section 2.3.1.1 : le nombre minimal d'occurrences).

La carte, présentée figure 8.7 [Schiefenhövel, 2014], a été créée sur la base de la carte de Ross *et al.* [2005] et modifiée selon des données récoltées sur les sites : « Ethnologue » (<http://www.ethnologue.com/>) et « Muturzikin » (<http://www.muturzikin.com/>). Cette carte

11. pour *Ecolinguistic Niche*

présente les aires qui sont attribuées à chacun des groupes linguistiques, nous appelons cette surface « aire linguistique ».

Nous avons collecté les occurrences selon la méthode décrite section 3.2 à partir des fichiers listant tous les villages d'Indonésie et de Papouasie (téléchargés sur le site DIVA-GIS¹²). Les affiliations linguistiques des villages ont été établies selon la carte présentée figure 8.7.

Ainsi, cette requête spatiale nous a permis de prélever des villages de la liste correspondant à chaque aire linguistique de la carte. Nous avons fixé la limite d'échantillonnage à 50 pour assurer une homogénéité du nombre d'occurrences par groupe linguistique. Un total de 1532 occurrences ont été prélevées, leurs localisations sont présentées figure 8.8.

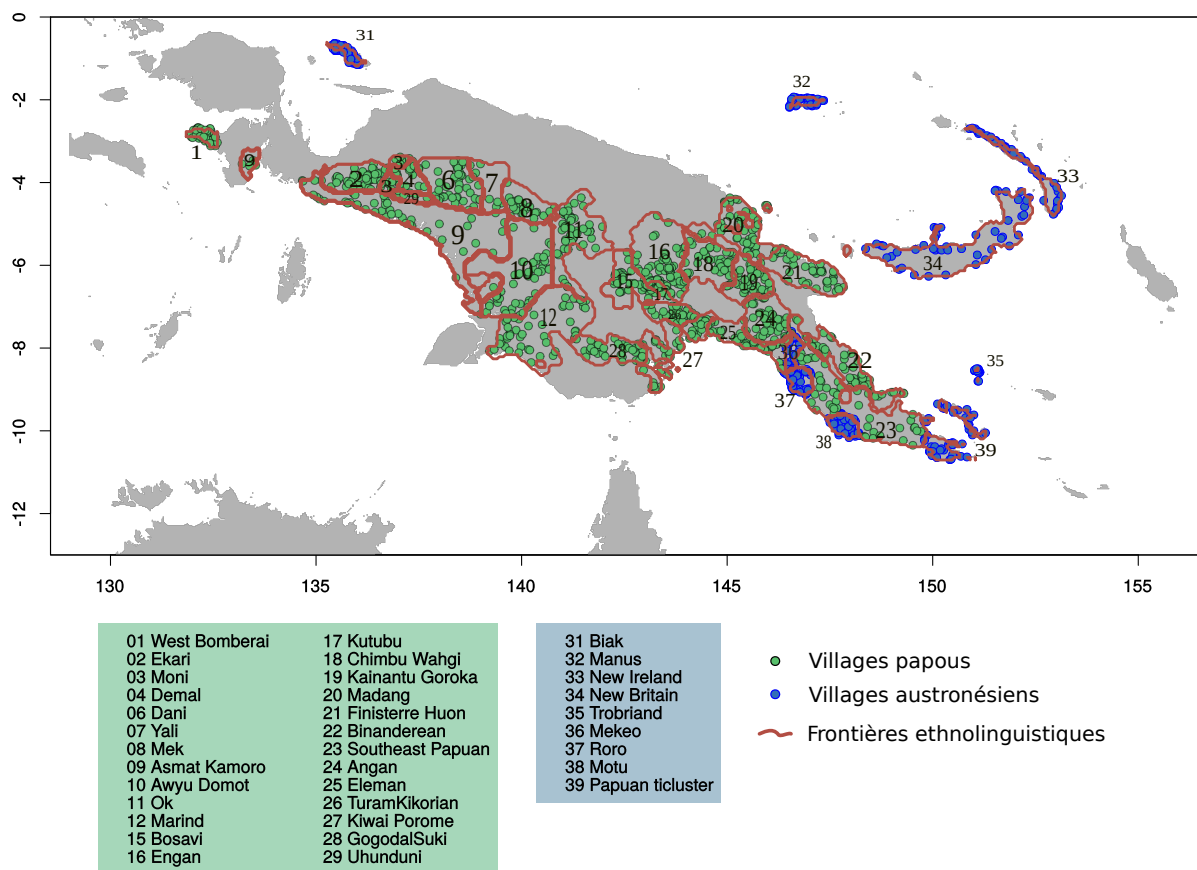


FIGURE 8.8 : Occurrences utilisées pour l'ELNM en Nouvelle-Guinée. Les points représentent les villages utilisés comme occurrences de présence pour l'ELNM (n=1532). Les numéros et les zones délimitées par des polygones indiquent les aires linguistiques des groupes de langues papoues (01-29) et austronésiennes (31-39).

12. <http://www.diva-gis.org/>

Afin d'alléger nos propos, nous ferons dans le reste du document, à plusieurs reprises, directement références aux locuteurs « Papous » ou « Austronésiens » au lieu de parler de groupes linguistiques de la famille papoue ou austronésienne.

8.5.2 Données environnementales

Nous avons choisi de modéliser les ELN de Nouvelle-Guinée à partir des 19 variables bioclimatiques (décrites tableau 2.6) provenant de la simulation *WorldClim* [Hijmans *et al.*, 2005] et des sept variables topographiques décrites tableau 2.5. Comme décrit dans la section 8.2.1, la Nouvelle-Guinée possède des paysages montagneux qui peuvent être très encaissés avec de nombreux canyons, pics et vallées pouvant créer des variations climatiques très locales. Ainsi, dans le but d'assurer une sensibilité micro-climatique aux prédictions de niches, nous avons mené l'ELNM avec la résolution maximale dont nous disposions, soit 30 arc-secondes ($\approx 1 \text{ Km}^2$).

8.5.3 Délimitation du M

Afin d'optimiser les prédictions de niches, nous restreignons l'aire d'étude selon une zone M (voir section 2.4.2.2). Cette aire a été définie relativement à la mobilité la plus restreinte des deux familles, c'est-à-dire celle de la famille de langues papoues, les Austronésiens étant de grands navigateurs (section 8.3.2). Nous avons donc considéré l'archipel entre les méridiens 130°E et 153°E et les parallèles 0° et 11°S. Lors de ce choix, nous avons exclu l'Australie présentant un environnement très différent qui pouvait biaiser l'analyse. Afin d'éviter le même biais, nous avons également écarté de l'analyse l'île de Timor (bien que cette île soit peuplée de Papous et d'Austronésiens), car elle se trouve à une distance importante de la Nouvelle-Guinée, ainsi que du reste du monde austronésien.

Le choix de cette large zone d'étude, malgré la territorialité restreinte de certaines langues, est justifié par la large capacité de diffusion géographique des groupes linguistiques modélisés dans cette étude. Effectivement, les Papous sont présents dans plusieurs îles de l'archipel avec une quasi-ubiquité au sein de l'île principale (où ils peuvent parcourir de longues distances pour acquérir des ressources [Weller *et al.*, 1996] et dans laquelle ils ne connaissent pas d'obstacle infranchissable [Schiefenhövel, 2014]). Les Austronésiens eux, possèdent des connaissances de navigation leur permettant d'atteindre toutes les îles de l'archipel étudié. D'autre part, nous devons garder à l'esprit que si la circulation des individus n'est pas fluide, filtrante ou impossible (pour des raisons pratiques ou résultant de relations sociales), certains éléments linguistiques comme le lexique, la syntaxe ou les expressions idiomatiques peuvent néanmoins franchir ces frontières [Chesterman, 1997; Heylighen, 1998].

8.5.4 Seuil d'omission

Établir un seuil d'omission en modélisation de niches permet de pallier la présence d'occurrences non représentatives due aux erreurs d'échantillonnages, d'attribution ou bien à un événement historique particulier. Dans notre cas, les villages compris dans une aire linguistique parlent une langue du groupe qu'il leur est affilié, néanmoins on ne peut exclure le fait que quelques villageois frontaliers, du fait de leurs interactions avec les voisins, parlent peu la langue du groupe ou que cette dernière ne soit pas leur langue principale. D'autre part, des annexions ou des pertes de territoires récentes peuvent également inclure des villages d'un autre groupe linguistique que celui qui leur correspond géographiquement. Enfin, la liste DIVA-GIS, dans laquelle nous avons puisé les occurrences, recense tous les villages de Nouvelle-Guinée. Or, les constructions de certains peuvent être récentes et donc en désaccord avec le système adaptatif originel de leurs habitants, dû à l'adoption de pratiques qui ont trait au mode de vie occidental. Pour ces trois raisons, nous appliquons un seuil d'omission avec une valeur relativement élevée ($E = 10\%$) qui nous permet d'accorder plus de confiance aux prédictions. Ce seuil est appliqué selon la méthode détaillée section 2.5.2 à toutes les prédictions avant de construire les consensus.

8.5.5 Consensus élitistes pondérés

Pour reconstruire les ELN des 35 groupes linguistiques, nous avons utilisé 11 algorithmes dont certains avec des configurations différentes ramenant le nombre de techniques de modélisations à 15 au total. L'interprétation de chacune des 525 ELN ainsi générées relèverait de l'impossible. De plus, comme nous pouvons le constater dans la figure 8.9, la pertinence des prédictions n'est pas régulière pour chaque technique. Afin d'obtenir une prédiction unique pour chaque groupe linguistique, tenant compte des prédictions de chaque algorithme et de leurs pertinences respectives, nous avons opté pour la création de consensus élitistes pondérés. Les prédictions retenues sont celles dont l'*accuracy* est supérieure à 80 et l'AUC supérieure à 0.9, la pondération est équivalente au *partial ROC ratio* des prédictions de niches (voir section 6.4).

8.5.6 Résultats de l'ELNM

Les projections géographiques des niches écolinguistiques papoues et austronésiennes sont représentées respectivement figures 8.10 et 8.11. Chaque sous-figure est composée de deux parties. Dans la partie de droite est représentée la projection géographique de la niche avec une échelle de probabilité variant de 0 à 1 (du gris au noir en passant dans l'ordre par le vert, le jaune et le rouge). Sur la gauche 3 volets indiquant chacun des informations sur le groupe linguistique modélisé :

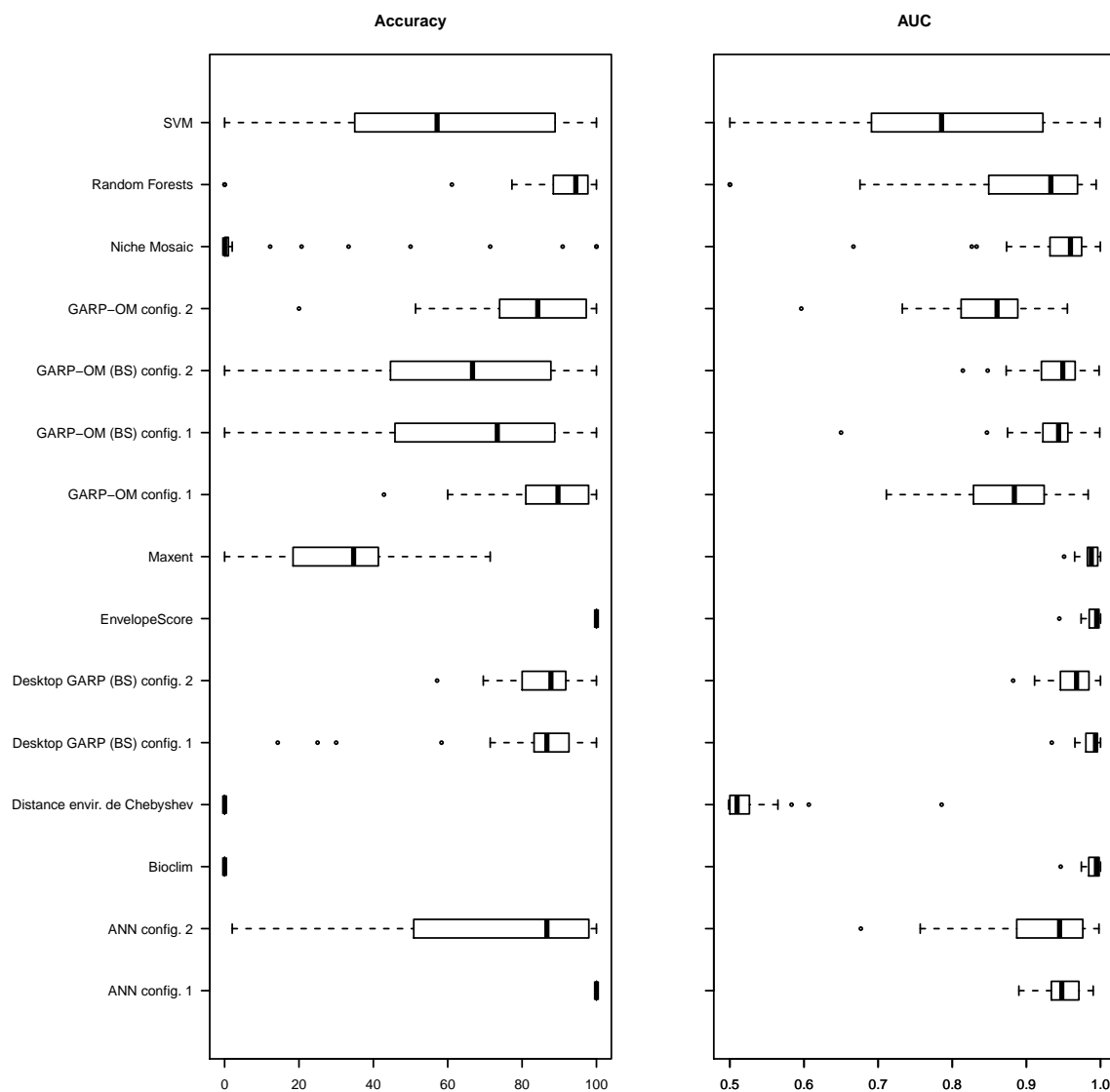
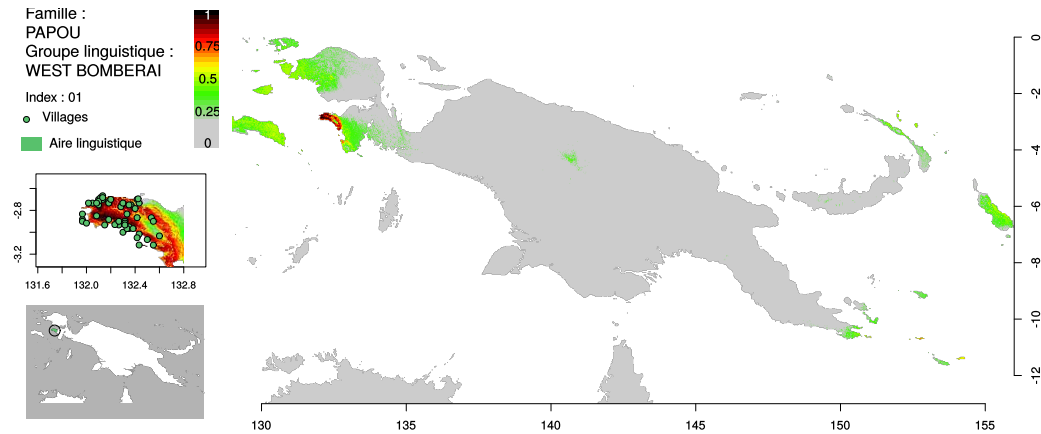
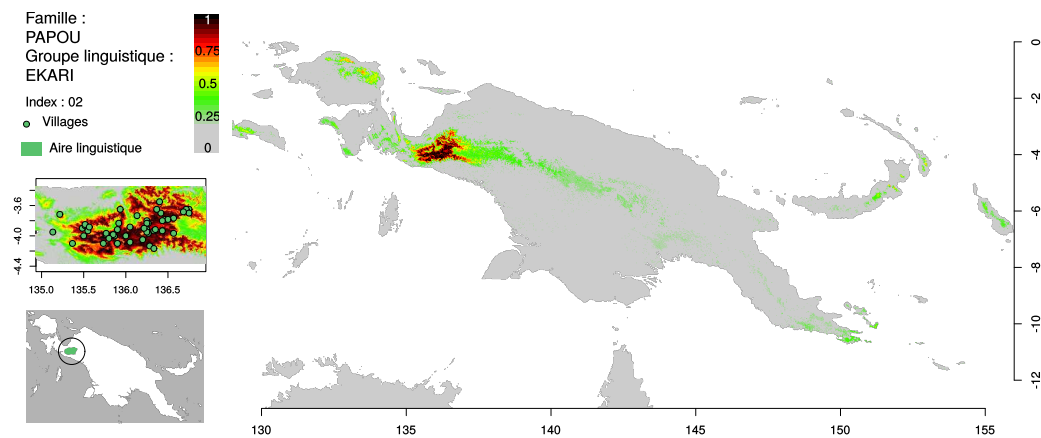


FIGURE 8.9 : Variations d'*accuracy* et d'AUC des algorithmes prédictifs pour les 525 reconstructions de niches écolinguistiques issues des 15 techniques de modélisations appliquées à chacun des 35 groupes linguistiques. *Open Modeller* a été utilisé pour réaliser la modélisation. Les lettres (BS) signalent l'utilisation de l'option *best subsets* de DesktopGARP et de la version GARP d'*Open Modeller* (GARP-OM). L'annotation : config 1-2 indique que certains algorithmes ont été utilisés avec plusieurs paramètres.

- un volet indiquant la famille linguistique à laquelle le groupe appartient, son nom et son index ¹³
- un zoom sur l'aire linguistique permettant d'observer la position de ses occurrences et les prédictions proches de ces dernières
- la localisation du groupe dans l'aire géographique étudiée.

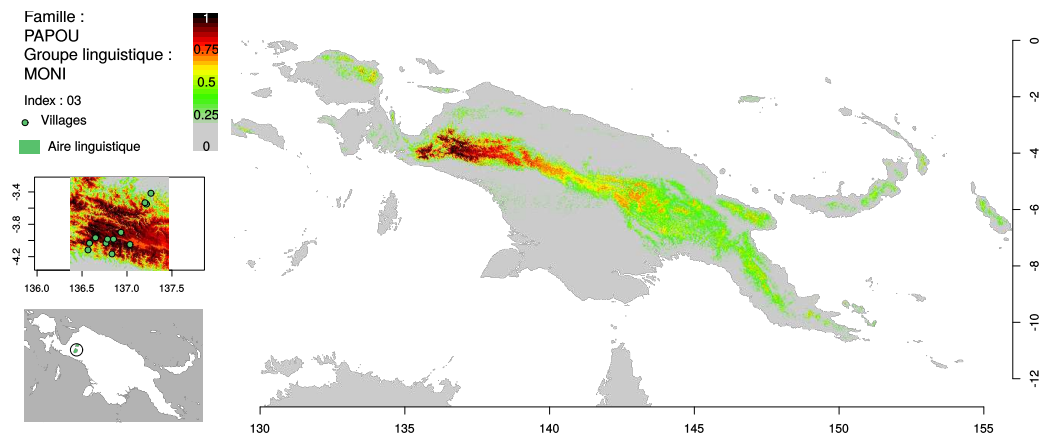


(a) West Bomberai

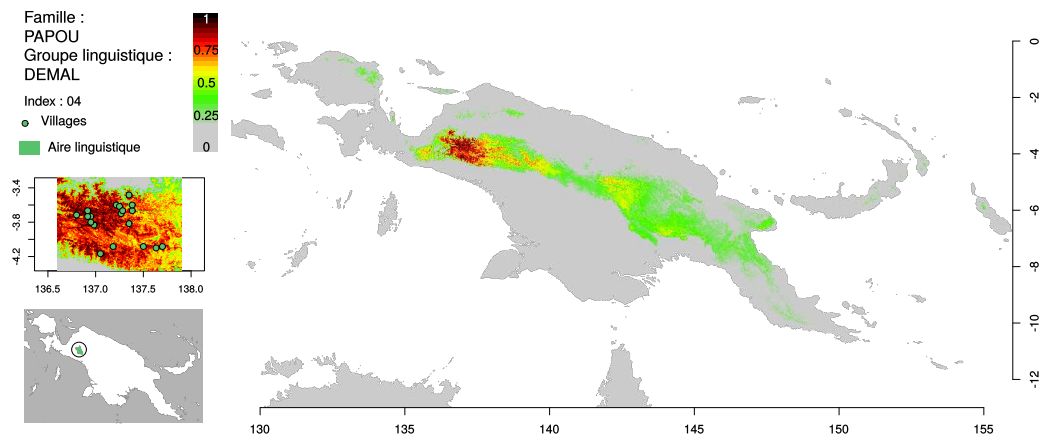


(b) Ekari

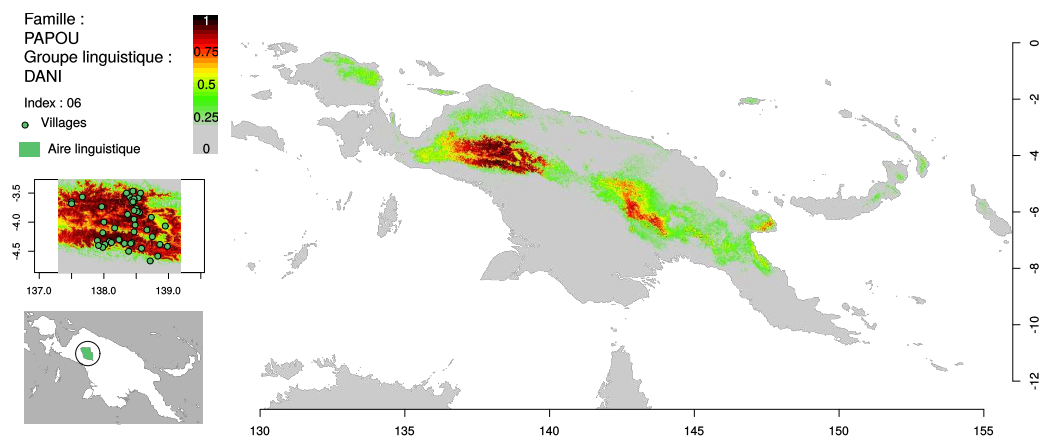
13. nous utilisons un index pour chaque groupe linguistique permettant de le situer sur les cartes 8.8 et 8.7, nous retrouverons ce même index dans les diverses figures de résultats ainsi que dans le texte de cette thèse



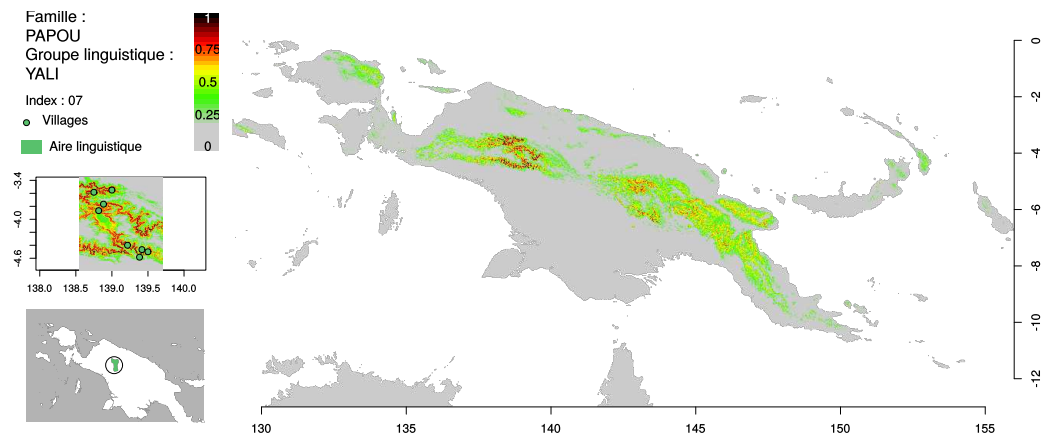
(c) Moni



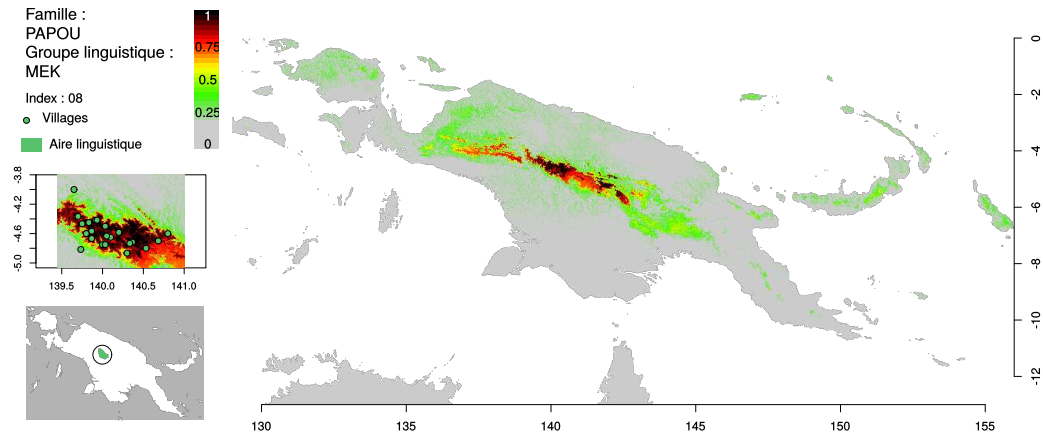
(d) Demal



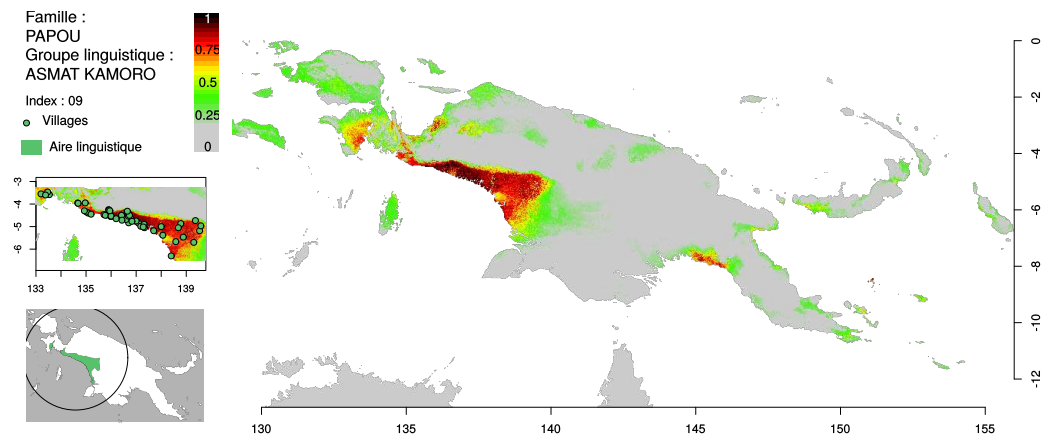
(e) Dani



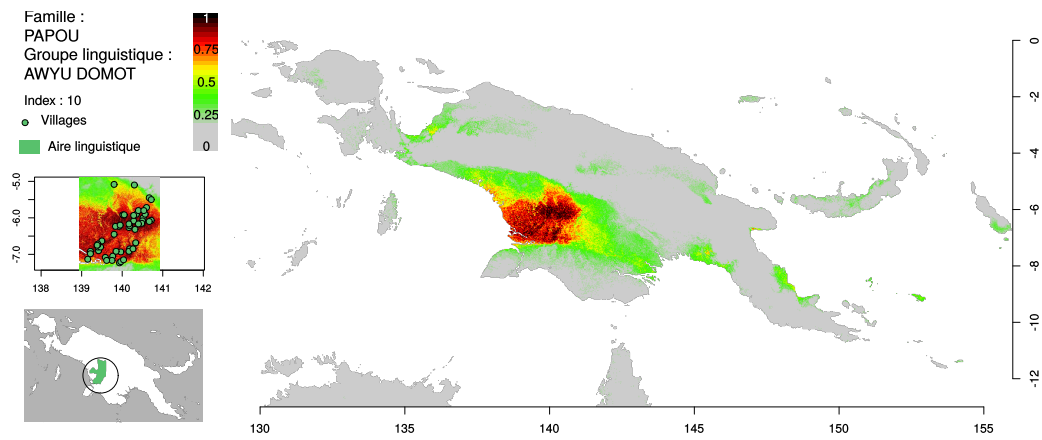
(f) Yali



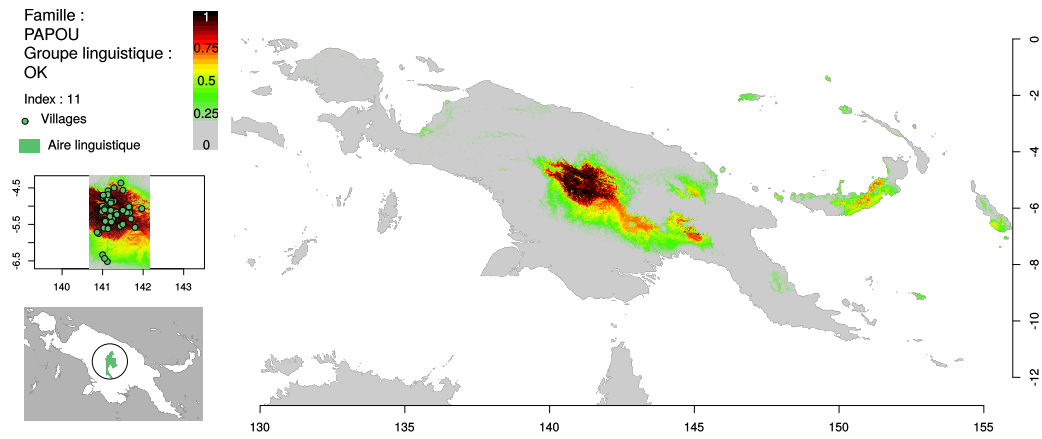
(g) Mek



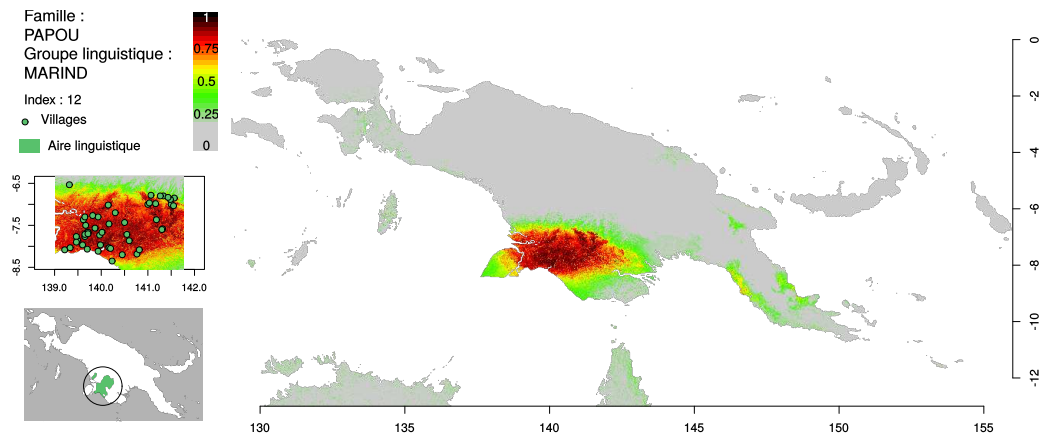
(h) Asmat-Kamoro



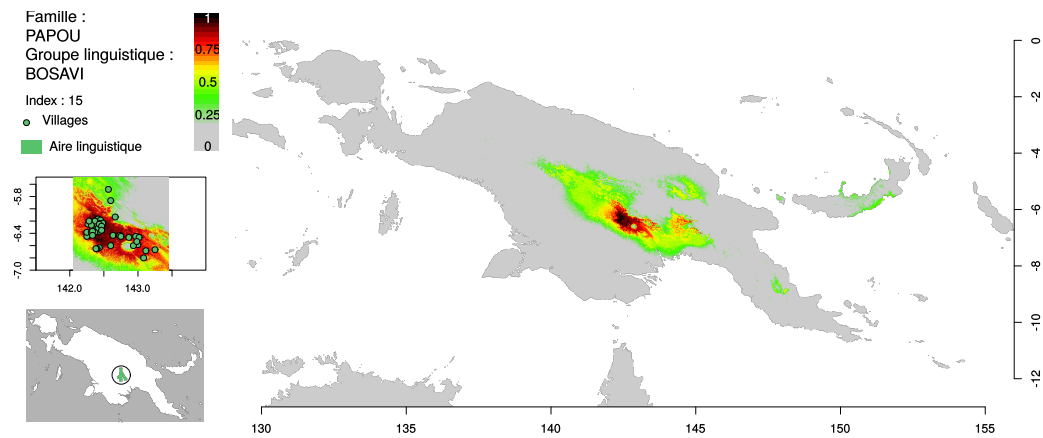
(i) Awyu-Domot



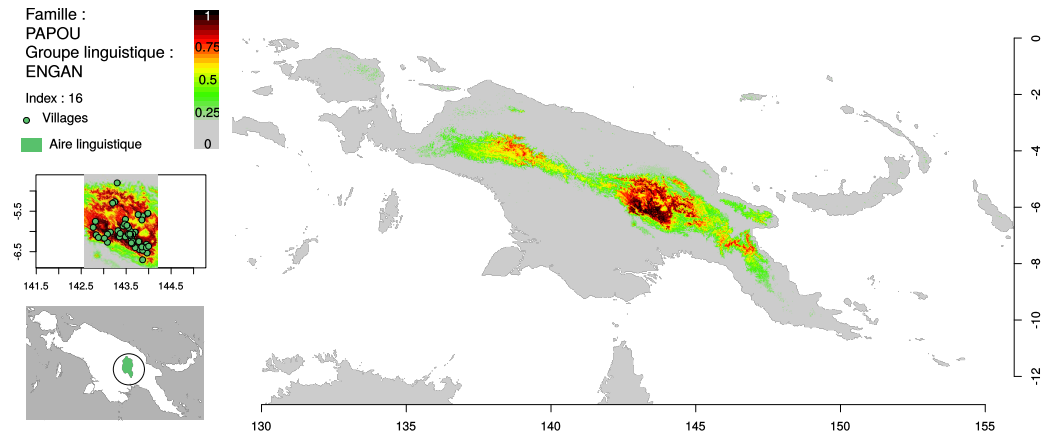
(j) Ok



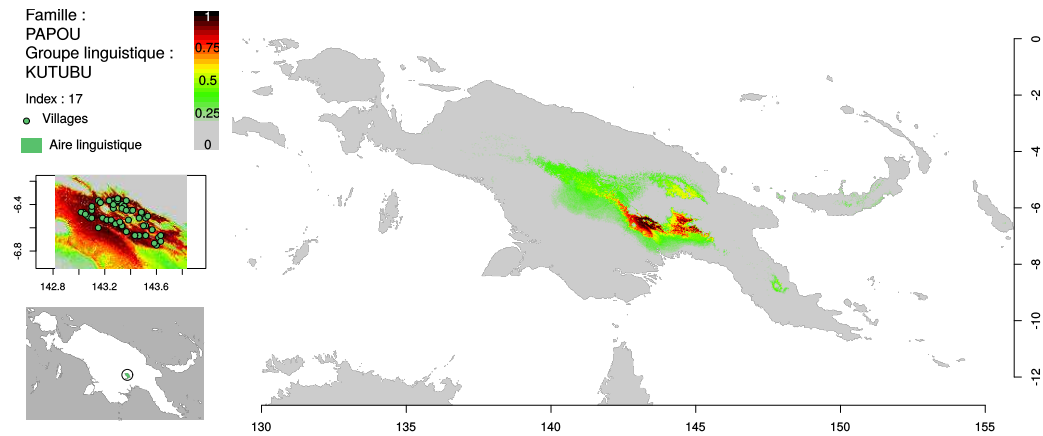
(k) Marind



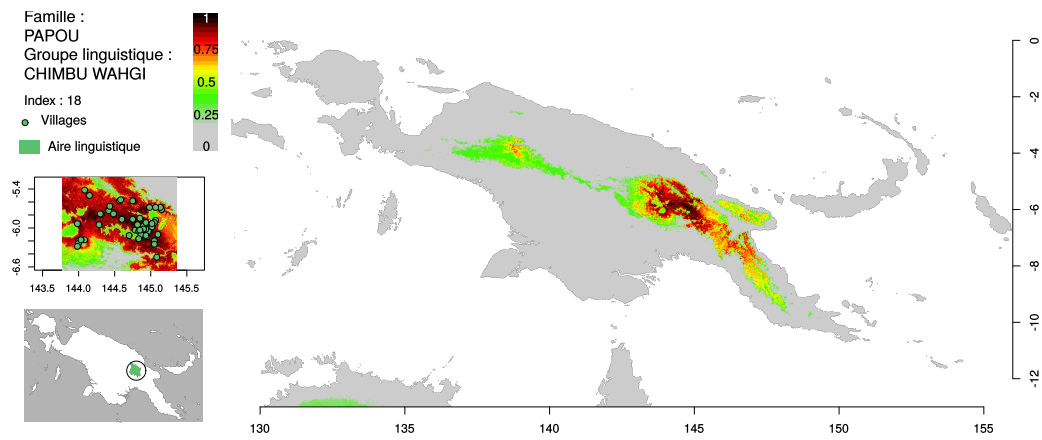
(l) Bosavi



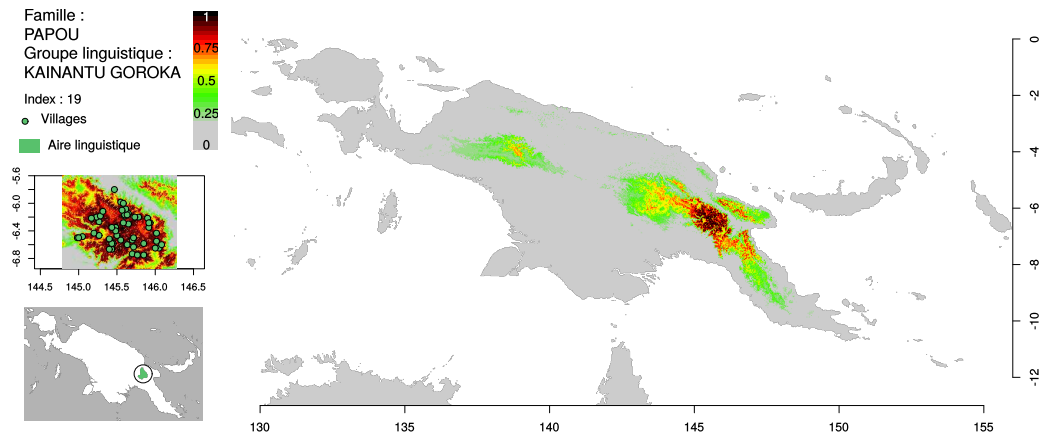
(m) Engan



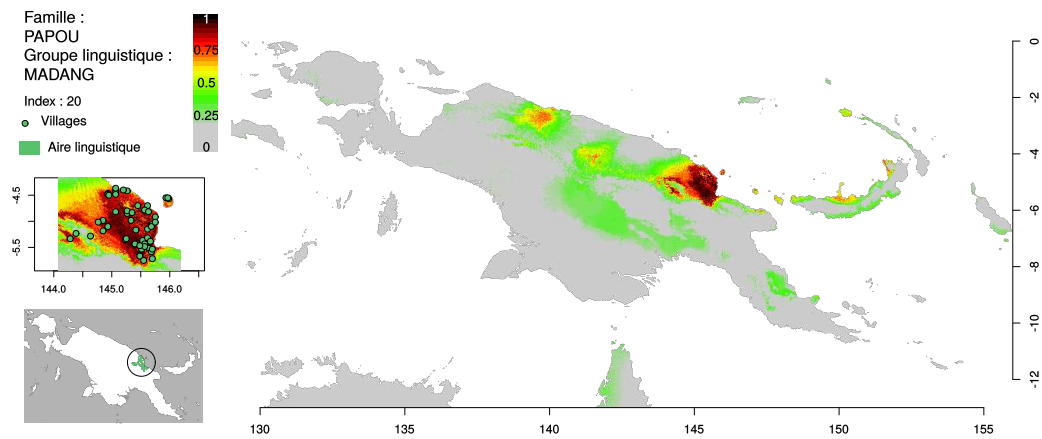
(n) Kutubu



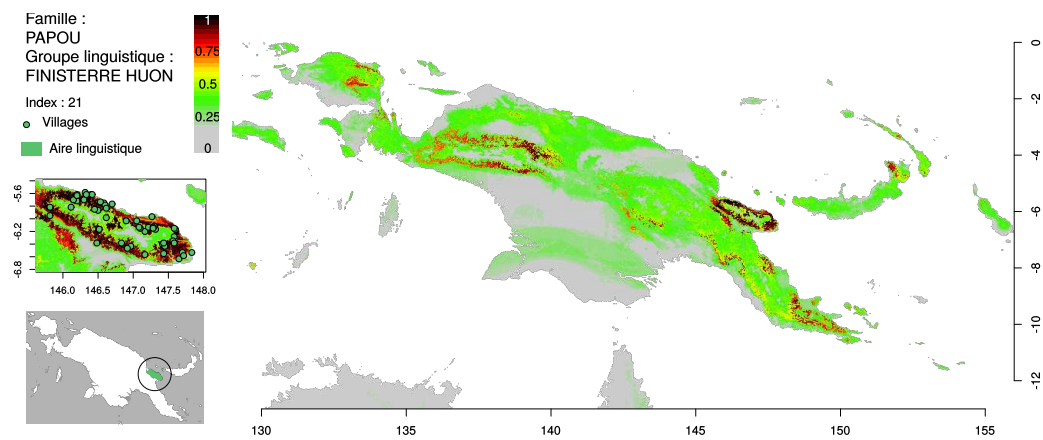
(o) Chimbu-Wahgi



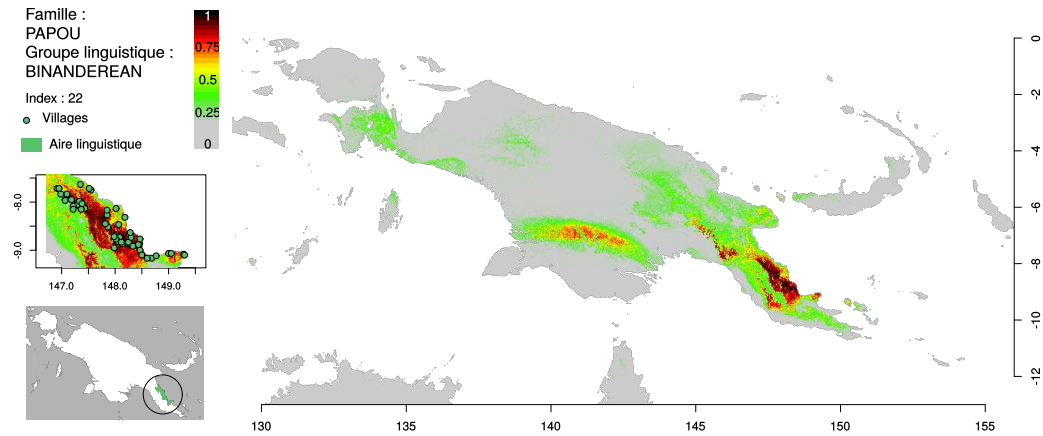
(p) Kainantu-Goroka



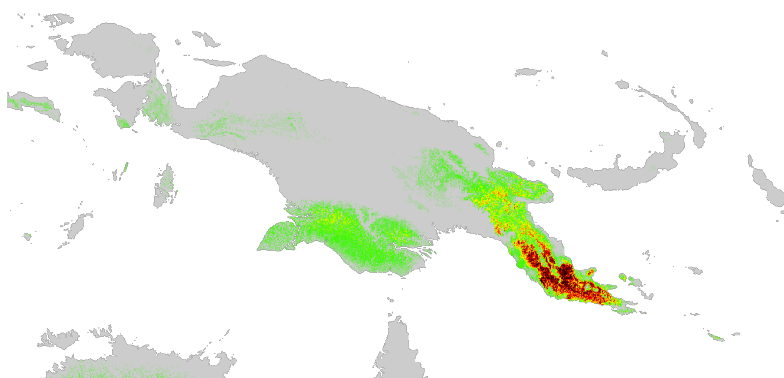
(q) Madang



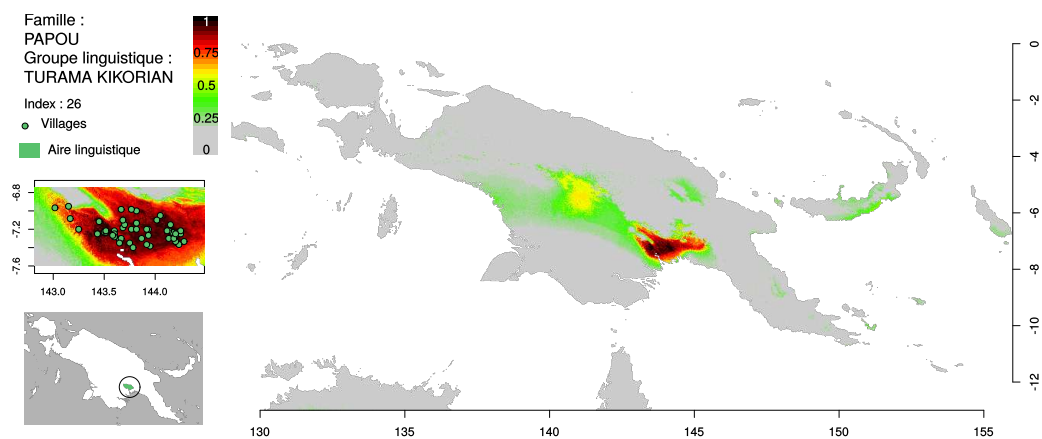
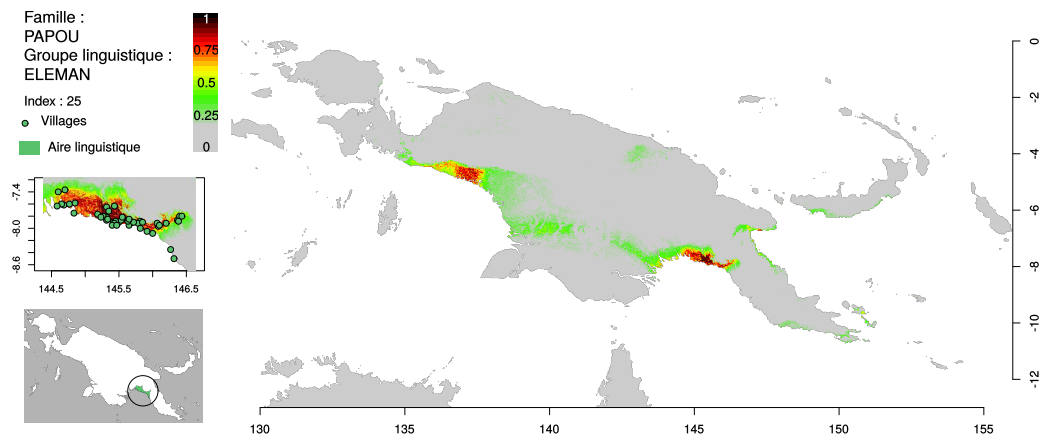
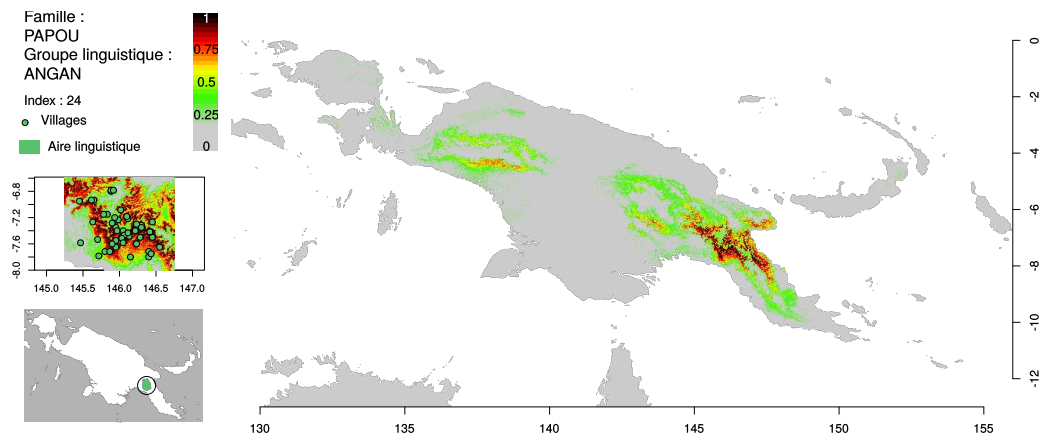
(r) Finisterre-Huon

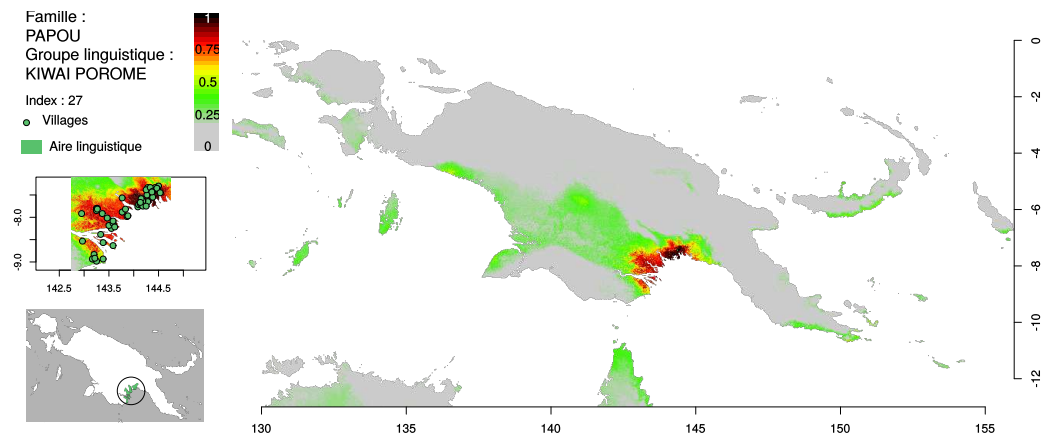


(s) Binanderean

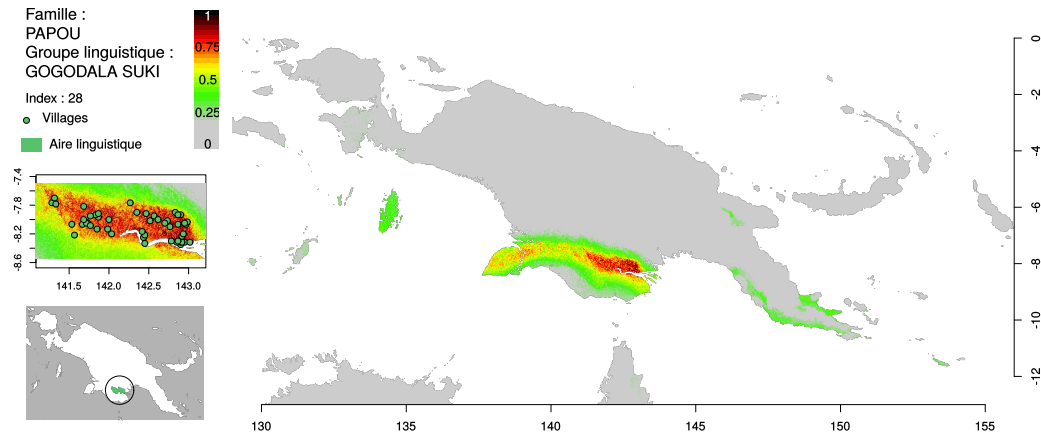


(t) Southeast-Papuan

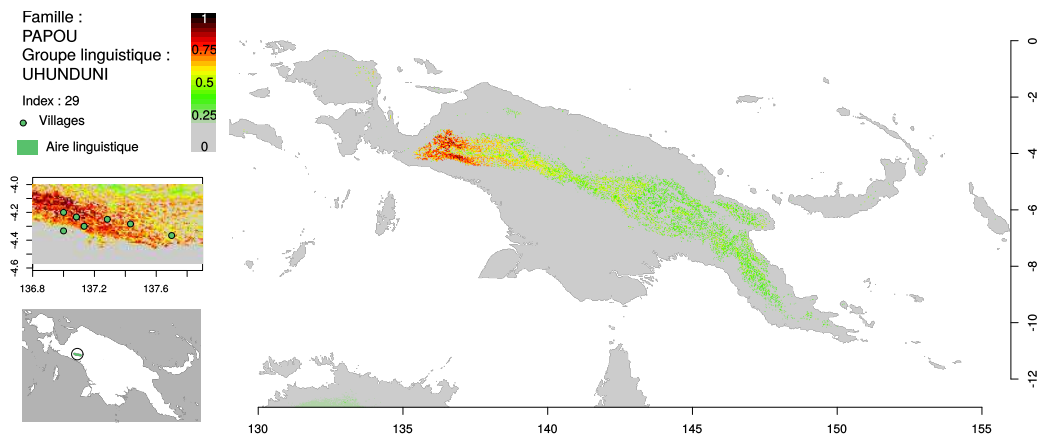




(x) Kiwai-Porome

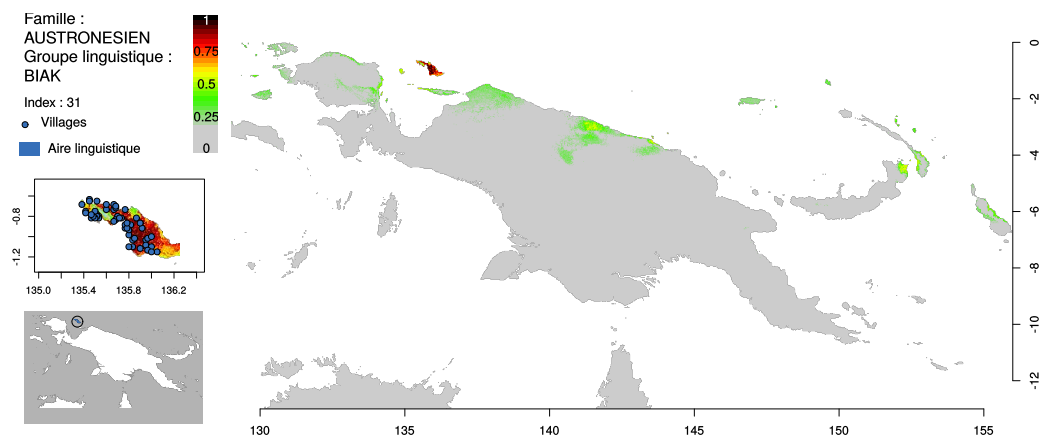


(y) Gogodala-Suki

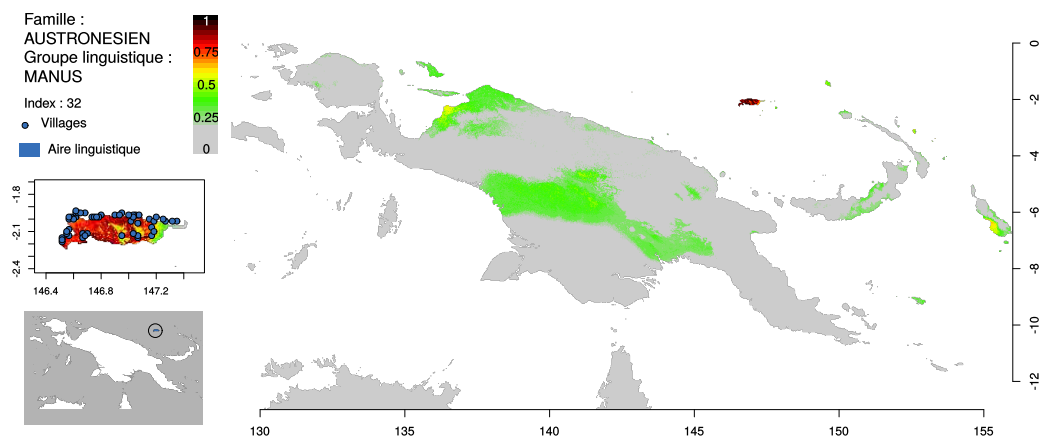


(z) Uhunduni

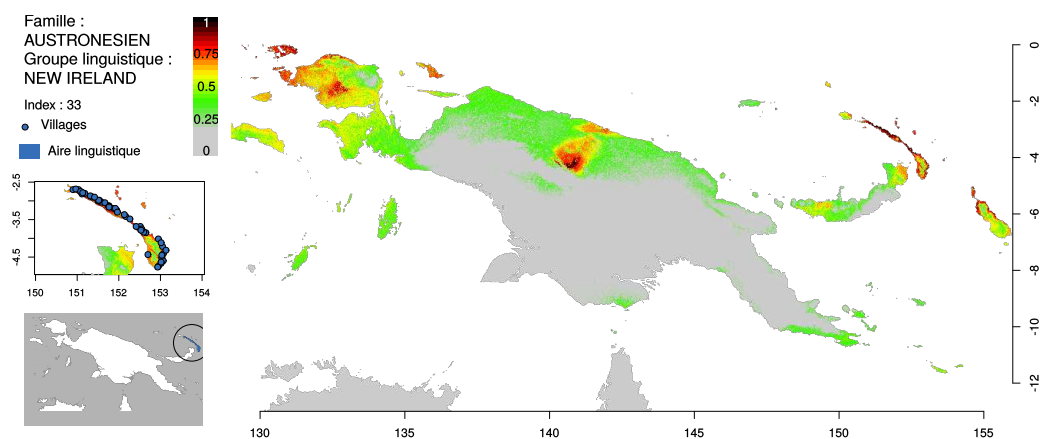
FIGURE 8.10 : Distributions géographiques des niches écolinguistiques de 26 groupes de langues papoues de Nouvelle-Guinée



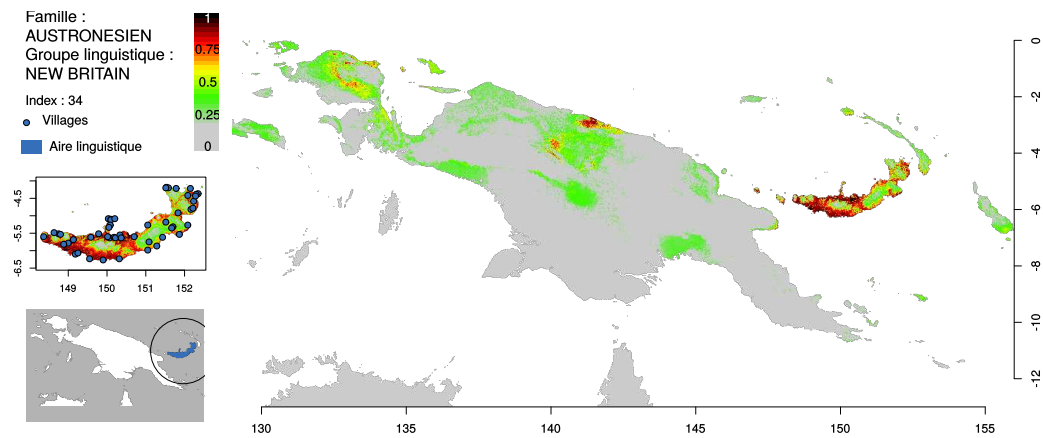
(a) Biak



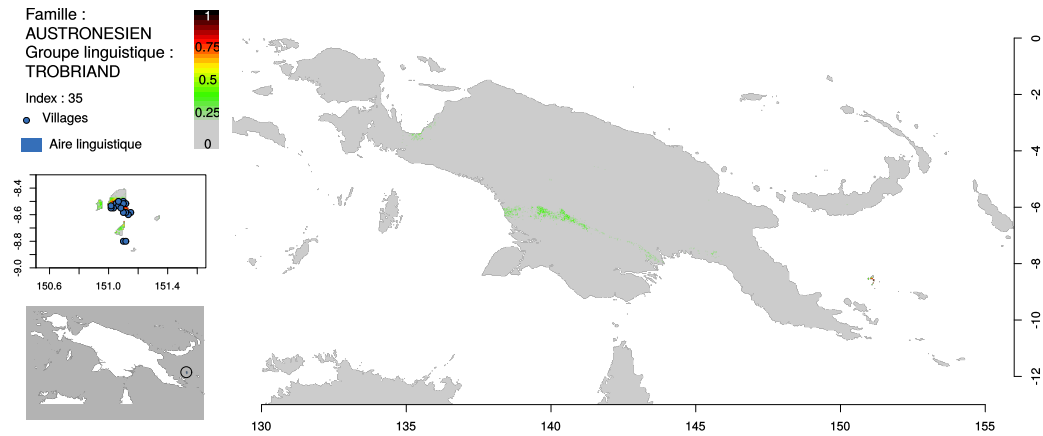
(b) Manus



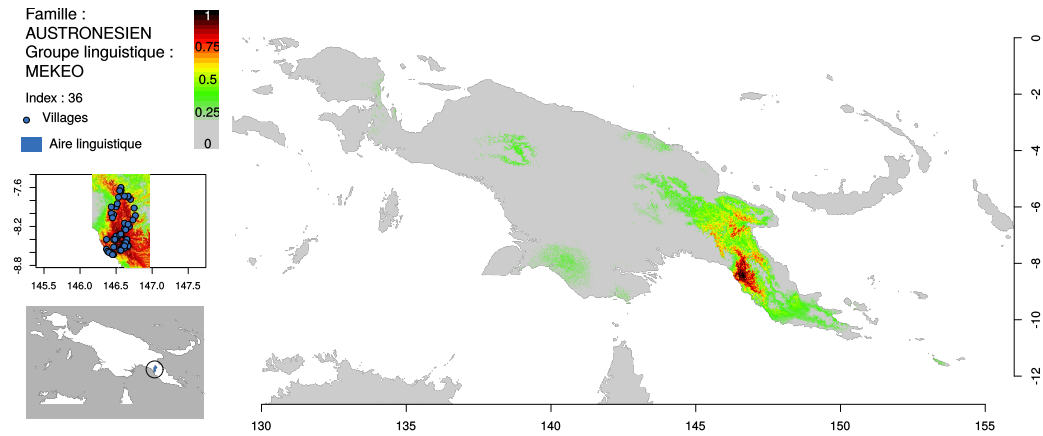
(c) New-Ireland



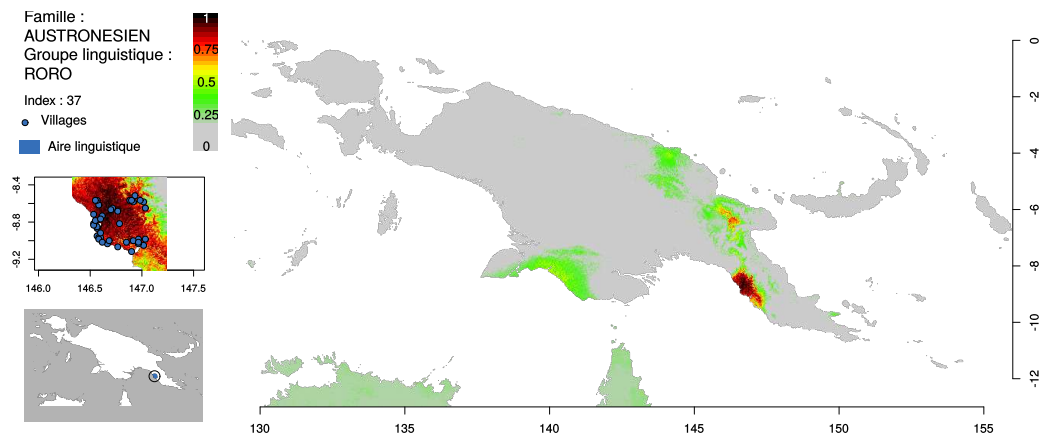
(d) New-Britain



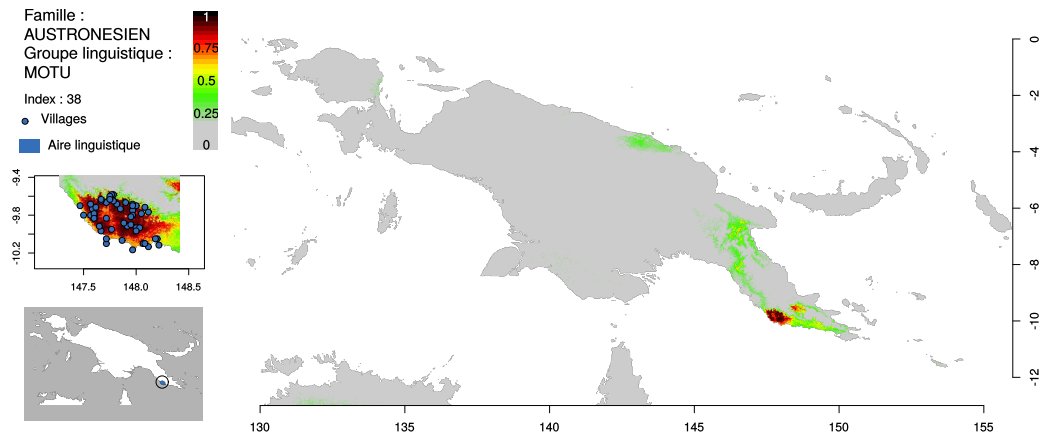
(e) Trobriand



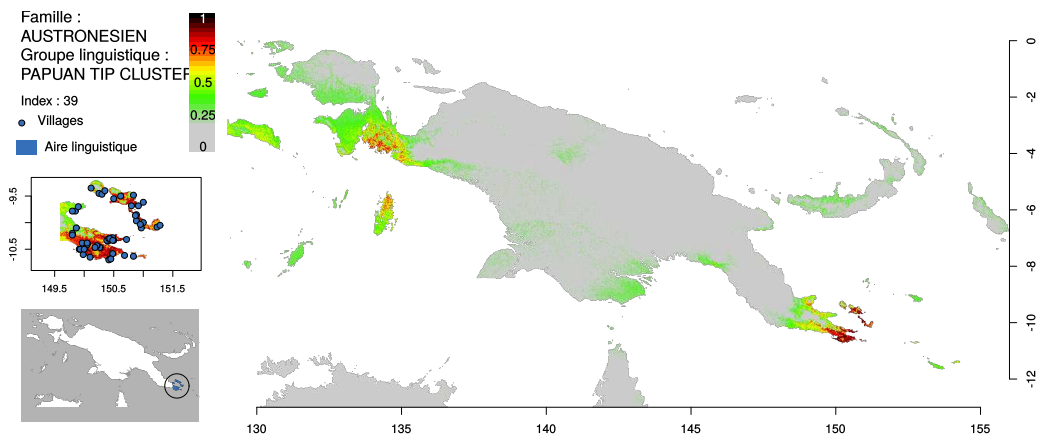
(f) Mekeo



(g) Roro



(h) Motu



(i) Papuan-tip-cluster

FIGURE 8.11 : Distributions géographiques de niches écolinguistiques de 9 groupes de langues austronésiennes de Nouvelle-Guinée.

Comme nous l'avons vu dans le chapitre 7, une analyse en composantes principales (ACP) permet d'apprécier les importances relatives des variables environnementales dans l'établissement des niches. Nous avons réalisé cette ACP pour les prédictions supérieures à 0.5.

8.5.6.1 Positions des ELN dans l'espace environnemental

L'ACP a été réalisée grâce au *package* : « *ade4* » [Dray *et al.*, 2007] du logiciel R [Team, 2013]. Les individus de l'ACP sont les localités prédites de chaque ELN pour lesquelles la probabilité de présence est supérieure à 0.5. Les variables de l'analyse sont les 26 variables environnementales. La figure 8.12 synthétise les résultats de cette analyse.

L'histogramme des valeurs propres, situé dans l'angle haut gauche du graphique, indique que les deux premiers axes suffisent à eux seuls pour expliquer la majeure partie des corrélations entre les variables. La pondération des axes est de 43.75 % pour le premier et 23.6 % pour le second. Une présentation des résultats dans le plan suivant l'axe 1 (axe des abscisses) et l'axe 2 (axe des ordonnées) est donc représentative de la position dans l'espace écologique des niches. Elle permet d'observer quelles sont les variables déterminantes pour chaque niche et de quantifier leur importance dans l'élaboration de ces dernières. Les individus de chaque groupe linguistique sont représentés par une ellipse tracée en fonction de l'inertie et dont le centre correspond à l'inertie moyenne de 61% des points. Les flèches représentent les variables environnementales.

En s'appuyant sur le diagramme figure 8.12, nous pouvons distinguer plusieurs regroupements entre groupes linguistiques influencés par les températures, les précipitations et la topographie. Nous divisons le plan de l'ACP en fonction des regroupements de niches que l'on peut effectuer relativement à leur proximité dans l'ACP et des formes de leurs expressions géographiques (figures 8.10 et 8.11). Nous qualifions de position écologique (p-eco) les régions du plan ainsi concernées. Les ELN des groupes linguistiques papous se répartissent sur l'ensemble du plan de l'ACP selon cinq positions écologiques :

- p-eco1** Les groupes linguistiques papous : 2, 3, 4, 6, 7, 8, 16, 18, 19 et 29, sont fortement corrélés aux variables topographiques (notamment à l'altitude) et plus faiblement liés à l'amplitude thermique. Ces groupes sont majoritairement répartis dans les Hautes Terres.
- p-eco2** L'ACP dévoile un deuxième regroupement (groupes linguistiques : 21, 22, 23, 24) au centre de l'ACP. Le groupe *Finisterre Huon* (21) a une niche très différente (figure 8.10(r)) mais sa position centrale dans le diagramme est due au fait que sa niche soit très large et hétérogène, recoupant presque l'ensemble des éco-régions de Nouvelle-Guinée. Les trois autres : 22, 23, et 24 constituent, selon leurs ressemblances écologiques, un groupe qui se situe dans le Sud-Est de l'île principale.
- p-eco3** Un troisième rassemblement d'ELN semble être dépendant des précipitations, il est composé des groupes linguistiques 11, 15, 17 et 26. Ces groupes linguistiques sont

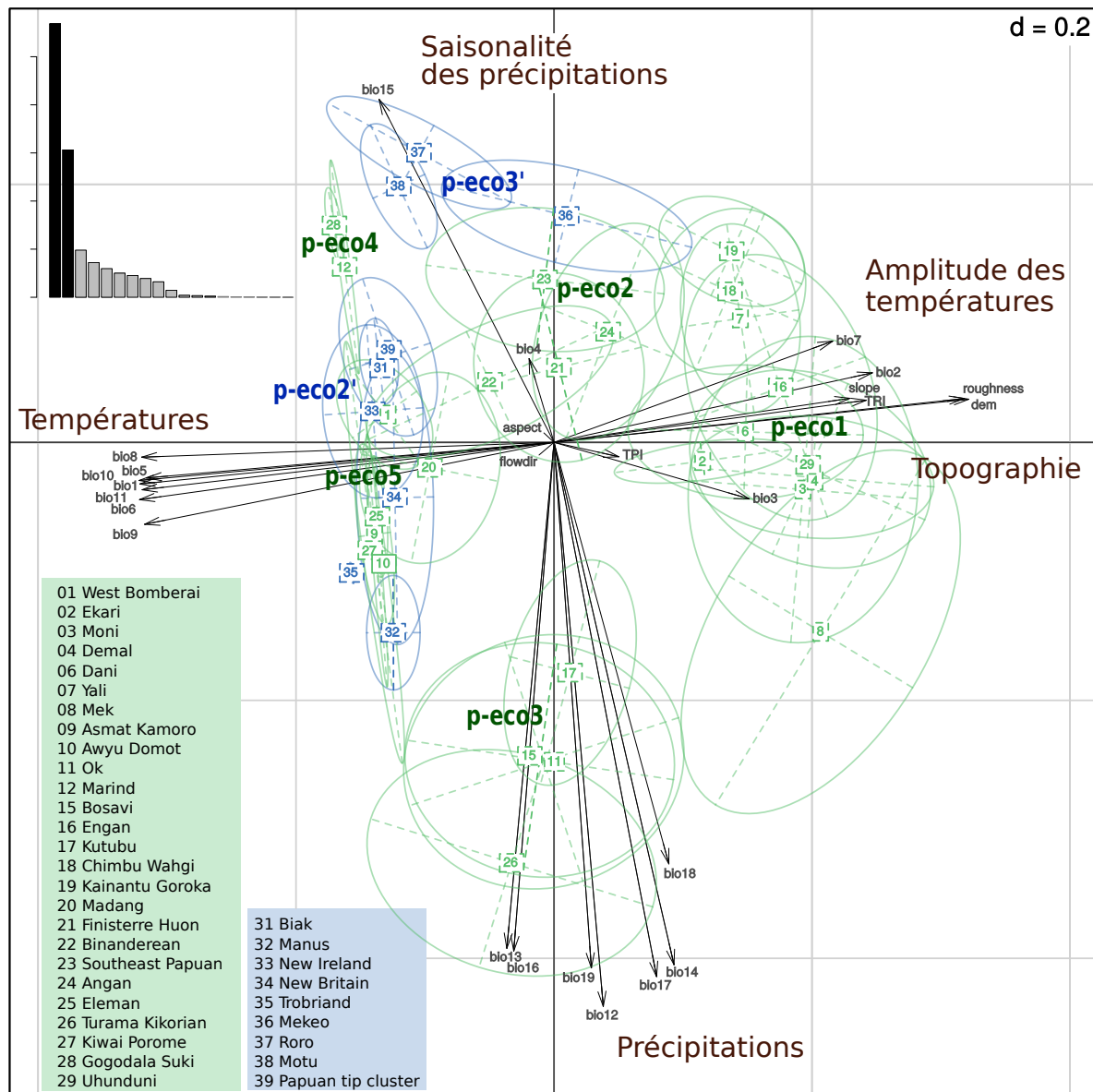


FIGURE 8.12 : Analyse en composante principale des groupes linguistiques papous et austronésiens de Nouvelle-Guinée. Dans le coin supérieur gauche se trouve l’histogramme des valeurs propres. L’axe 1 a une inertie de 43.75 % et l’axe 2 de 23.6 %. Les flèches représentent les variables bioclimatiques (tableau 2.6) et topographiques (tableau 2.5). Les positions des numéros indiquent les centres des ellipses d’inerties des prédictions des niches écolinguistiques. Les indications : saisonnalité des précipitations, températures, amplitude des températures, précipitations et topographie, synthétisent les informations apportées par les variables qui composent les axes. En vert : p-eco1-5 indiquent la région du plan de l’ACP pour les positions écologiques des Papous ; en bleu : p-eco1’-3’ indiquent les positions écologiques des Austronésiens.

situés au centre de la Nouvelle-Guinée excepté *Turama-kikorian* (26) qui se trouve dans le golfe de Papouasie. Lorsqu'on regarde l'expression géographique de la niche de ce groupe (figure 8.10(w)), on s'aperçoit alors que les prédictions sont scindées en deux : une partie avec des prédictions fortes dans la région de l'aire linguistique et une autre région, dont les probabilités de présence prédites sont plus faibles, proche des trois autres groupes (11, 15 et 16). Le groupe 26 est donc à exclure de cet ensemble écologique de niche car sa position dans le plan de l'ACP est biaisée par la valeur du seuil de probabilité de présence que nous avons fixé à 0.5.

p-eco4 28 et 12 sont des ELN très semblables que l'on retrouve au Sud des Hautes Terres entre la mer d'Arafura et le golfe de Papouasie. Elles sont principalement influencées par les températures et la saisonnalité des précipitations.

p-eco5 Les groupes de langues papoues 1, 9, 10, 20, 25 et 27 apparaissent dans l'ACP comme ayant des structures écologiques fortement dépendantes des températures et secondairement influencées par les précipitations, cependant cet ensemble ne présente pas de niches similaires géographiquement.

Hormis le groupe linguistique *Mekeo* (36), les Austronésiens ont environ la même valeur en première composante, la température et les basses altitudes sont donc les principaux facteurs liés à l'élaboration de leurs niches. L'axe 2 permet, quant à lui, d'identifier les facteurs environnementaux qui rentrent en jeu dans la différenciation des ELN de ces groupes linguistiques. Trois positions écologiques de niches écolinguistiques austronésiennes peuvent être distinguées. Ces groupes se situent sur un gradient relatif à l'intensité et aux fréquences des précipitations :

p-eco1' 32, 34, 35 est le groupement le plus humide. Il s'agit des ELN de trois îles situées à l'Est de l'île centrale.

p-eco3' 37, 38 (et 36, se trouvant légèrement à l'écart car le seuil fixé englobe une région prédite à altitude plus élevée), constituent le groupe d'ELN pour lequel la saisonnalité des précipitations est la plus marquée. Les trois groupes linguistiques se situent dans le Sud-Est de la Papouasie.

Le choix de ne considérer dans l'ACP que les prédictions supérieures à 0.5 permet d'apprécier la variabilité au sein de chacune des niches et de faire ainsi des regroupements entre les ELN selon l'importance des variables environnementales qui rentrent dans leur composition. Toutefois, le fait d'avoir fixé cette valeur de sélection basse, rapproche certains groupes de telle sorte qu'il est difficile d'observer les nuances entre leurs positions dans l'espace écologique. Regarder en détail les formes de ces niches, leurs superpositions et les écosystèmes qu'elles englobent permet de pallier ce biais.

8.5.6.2 Formes des d'ELN : distinction de deux catégories

Il est possible de classer les prédictions de niches écolinguistiques en deux catégories selon si la prédiction ou l'aire linguistique ont une surface plus ou moins équivalente (figure 8.13).

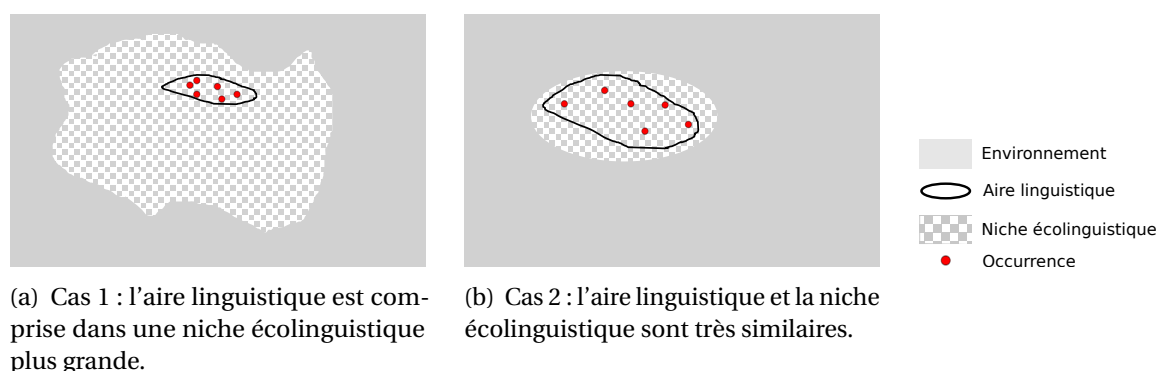


FIGURE 8.13 : Comparaison de l'aire linguistique et de la niche écolinguistique de populations théoriques.

Cas 1 : le groupe linguistique n'occupe qu'une partie de sa niche. Dans ce cas, la prédiction est maximale pour tout ou pour la majeure partie de l'aire linguistique et étend sa surface avec une forte probabilité à une zone plus large pouvant englober plusieurs éco-régions. Il peut subvenir toutefois quelques ambiguïtés dans l'affectation de cette catégorie à une ELN. Effectivement, un tel cas se produit lorsque la projection géographique de la niche dépasse largement l'aire linguistique, mais que les prédictions en dehors de celle-ci sont relativement faibles. Dans cette situation, on classe alors l'ELN dans la catégorie 1*.

Nous avons relevé cinq formes géographiques de niches partagées par les ELN de type 1 ou 1* de plusieurs groupes linguistiques. Nous qualifions ces formes de « motifs écologiques »¹⁴, lesquels peuvent être mis en relation avec les éco-régions (figure 8.4) :

- Les niches des Hautes Terres : cette forme est la plus commune, elle compte dix groupes linguistiques (2, 3, 4, 6, 7, 8, 16, 19, 18, 29) dont les prédictions de niches écolinguistiques s'étendent tout au long de la cordillère centrale de la Nouvelle-Guinée et possède des écosystèmes de forêts équatoriales de montagnes et de prairies alpines. Ce motif est parfaitement identifié par l'ACP (section 8.5.6.1) et correspond à la position écologique p-ecol.
- Les niches du Centre : motif constitué de trois groupes linguistiques (11, 15, 17) ayant une ELN prédisant une forte probabilité de présence au centre de la Nouvelle-Guinée,

14. nous donnons des noms propres à chacun de ces motifs en accord avec leur localité ou au paysage dominant.

c'est une zone de moyenne montagne où les écosystèmes dominants sont ceux de la forêt équatoriale de montagne. L'ACP situe ces groupes en position p-eco3. Notons que l'ELN des *Ok* (figure 8.10(j)) couvre en grande partie ce motif écologique bien que certains villages soient situés en altitude aussi importante que ceux des Hautes Terres (Kiwiron à 1800 m et Apmissibil 2500 m).

- Les niches des Marécages : ce motif réunit des groupes linguistiques du Sud. Les écosystèmes y sont principalement des forêts marécageuses d'eau douce et des forêts équatoriales de basse altitude. Sont concernés deux groupes linguistiques en position p-eco4 auxquels on peut rajouter les *Auyu-Domots* (10) qui se trouvent à l'écart dans l'ACP, car l'ELN de ce groupe englobe aussi des éco-régions à mangroves.
- Les niches des Mangroves : deux groupes linguistiques (26 et 27) offrent des motifs d'ELN très proches dans les régions à mangroves du golfe de Papouasie. Ces groupes ne sont pas classés par l'ACP dans une position particulière.
- Les niches du Sud-Est : ce regroupement est constitué de trois niches (22, 23, 24) qui ont une partie commune dans les Monts Owen Stanley (forêt équatoriale), cependant ces groupes linguistiques débordent légèrement sur d'autres éco-régions comme la savane, la forêt marécageuse et la forêt équatoriale de montagne. Cette observation correspond à la position p-eco2 et confirme l'exclusion du groupe de langues *Finisterre Huon* (21) par ailleurs situé dans sa propre éco-région.

Cas 2 : Le groupe linguistique et la niche prédite se superposent et leurs surfaces sont quasiment équivalentes. Ce type d'ELN est rencontré dans trois cas de façon stricte et dans 8 cas de façon moins tranchée. En effet, si la niche est scindée en plusieurs parties relativement éloignées dont une se superpose presque parfaitement avec l'aire linguistique on notera ce cas : 2*. Par définition, les ELN de type 2 et 2* ne peuvent donc pas être regroupées en motifs écologiques (puisque leur forme est unique).

Remarquons que si la mobilité d'une population est très élevée, alors sa niche qualifiée de 2* peut être en réalité considérée comme une niche de type 1. En effet, dans une telle situation, les milieux éloignés du groupe linguistique modélisé prédits comme propices par l'ELNM ne peuvent plus être considérés comme des îlots inaccessibles par les locuteurs. Nous constatons que l'attribution de classes d'ELN de type 1, 1*, 2 ou 2* est quelque peu subjective et peut être peu évidente dans certains cas.

En recoupant les informations apportées par cette classification, les données de l'ACP (figure 8.12) et les représentations géographiques des niches (figures 8.10) et 8.11, il est possible de corréler les ELN des groupes linguistiques avec la répartition des écosystèmes (figure 8.4) et dresser le tableau 8.1.

Nous allons voir qu'il est possible de relier les catégories de niches à l'intensité d'exploitation du milieu et donc au risque écologique. L'augmentation du risque écologique

Tableau 8.1 : Regroupements des ELN en fonction de leur projections géographiques, des éco-systèmes qu'elles recouvrent et de leur catégorie associée relativement à leur forme. La colonne éco-régions n'indique que les principales éco-régions recouvertes par les prédictions des ELN. Les lignes hachurées délimitent des différents motifs écologiques, les lignes pointillées indiquent des différences au niveau des écorégions. Cette table a été dressée à la fois en fonction de l'ACP effectuée sur les prédictions et sur l'analyse des expressions géographique des niches.

Motif	Index	Groupe de langues	Famille	Cas	Position écologique	Éco-régions
Hautes Terres	2	<i>Ekari</i>	Papoue	1*	p-eco1	Forêts équatoriales de montagne et prairies alpines
	3	<i>Moni</i>	Papoue	1	p-eco1	
	4	<i>Demal</i>	Papoue	1	p-eco1	
	6	<i>Dani</i>	Papoue	1	p-eco1	
	7	<i>Yali</i>	Papoue	1	p-eco1	
	8	<i>Mek</i>	Papoue	1	p-eco1	
	16	<i>Engan</i>	Papoue	1	p-eco1	
	19	<i>Chimbu-Wahgi</i>	Papoue	1	p-eco1	
	18	<i>Kainantu-Goroka</i>	Papoue	1	p-eco1	
	29	<i>Uhundini</i>	Papoue	1	p-eco1	
Sud-Est	22	<i>Binanderean</i>	Papoue	1	p-eco2	Forêts équatoriales et forêts marécageuses d'eau douce
	23	<i>Southeast-Papuan</i>	Papoue	1	p-eco2	Forêts équatoriales et savanes
	24	<i>Angan</i>	Papoue	1	p-eco2	Forêts équatoriales de montagne et du Sud-Est
Centre	11	<i>Ok</i>	Papoue	1	p-eco3	Forêts équatoriales de montagne
	17	<i>Kutubu</i>	Papoue	1	p-eco3	
	15	<i>Bosavi</i>	Papoue	1	p-eco3	
Marécages	12	<i>Marind</i>	Papoue	1	p-eco4	Forêts équatoriales de basse altitude
	28	<i>Gogodala-Suki</i>	Papoue	1	p-eco4	(+ mangroves)
	10	<i>Awyu-Domot</i>	Papoue	1	p-eco5	
Mangroves	20	<i>Madang</i>	Papoue	2*	p-eco5	Forêts marécageuses d'eau douce et forêts équatoriales de basses altitudes
	26	<i>Turama-Kikorian</i>	Papoue	1*	—	Mangroves, forêts marécageuses d'eau douce, et forêts équatoriales de basse altitude
	27	<i>Kiwai-Porome</i>	Papoue	1*	p-eco5	
	9	<i>Asmat-kamoro</i>	Papoue	2*	p-eco5	Mangroves et forêts marécageuses d'eau douce
	25	<i>Eleman</i>	Papoue	2*	p-eco5	Mangroves et forêts marécageuses d'eau douce
	1	<i>West Bomberai</i>	Papoue	1*	p-eco5	Forêts équatoriales de basses altitudes
	21	<i>Finisterre-Huon</i>	Papoue	1	—	(Pas de relations nettes car les prédictions sont très étalées)
	31	<i>Biak</i>	Austro.	2	p-eco2'	Forêt équatoriale
	33	<i>New Ireland</i>	Austro.	2*	p-eco2'	Forêts équatoriales de basses altitudes et forêts marécageuses d'eau douce
	34	<i>New Britain</i>	Austro.	2*	p-eco1'	Forêts équatoriales de basses altitudes et forêts marécageuses d'eau douce
	32	<i>Manus</i>	Austro.	2	p-eco1'	Forêts équatoriales
	35	<i>Trobriand</i>	Austro.	2	p-eco1'	forêts équatoriales
	36	<i>Mekeo</i>	Austro.	1	p-eco3'	Forêts équatoriales du Sud-Est et forêts équatoriales de montagne
	37	<i>Roro</i>	Austro.	2*	p-eco3'	Forêts équatoriales du Sud-Est, forêts marécageuses d'eau douce et savanes
	38	<i>Motu</i>	Austro.	2*	p-eco3'	Forêts équatoriales du Sud-Est
	39	<i>Papuan tip cluster</i>	Austro.	2*	p-eco2'	Forêts équatoriales du Sud-Est et forêts de basses altitudes

contraint les groupes humains à étendre leur territoire pour accéder à la quantité de ressources nécessaires à leur subsistance (voir section 1.6.4). Cet étalement est par ailleurs accompagné d'une cohésion sociale selon l'hypothèse de [Nettle \[1998\]](#). En ce sens, la classe 2 des ELN, qui correspond à l'extension maximale d'un groupe linguistique dans sa niche accessible, signalerait un risque écologique relativement élevé. À l'inverse, les niches classées 1 pourraient indiquer que les populations modélisées se trouvent dans un environnement plus clément. Ce milieu permettrait aux groupes de vivre en autarcie, dans ce cas puisque la cohésion sociale inter-groupes n'est plus essentielle à la survie du groupe originel, une augmentation de la diversité culturelle au sein du motif écologique pourrait se manifester par le biais de dérives comportementales.

Nous qualifions un environnement de « permissif » lorsqu'il offre la possibilité à des divergences culturelles ou fonctionnelles d'apparaître et de se maintenir. Les divergences induites provoquent la scission de la population en plusieurs groupes culturels. La permissivité des environnements envers les variations de systèmes adaptatifs peut donc être estimée en mesurant la diversité culturelle au sein de chaque motif écologique. De cette manière, on attribue aux régions prédites par les ELN de catégorie 2 une permissivité de 1 car elles ont par définition un motif écologique unique. Puis on attribue à chacun des motifs une permissivité écologique relative au nombre d'ELN qui les composent. Ainsi, le motif des Hautes Terres possède une permissivité égale à 10, celui du Centre une permissivité de 3, etc.

Nous venons de voir que la comparaison des projections géographiques des ELN avec les éco-régions permettait d'estimer la capacité d'accueil de l'environnement en termes de groupes linguistiques. Cependant, les informations concernant les écosystèmes ne sont pas toujours disponibles, notamment dans les cas d'études concernant des populations du passé. C'est pourquoi nous avons développé une autre approche basée sur l'ELNM, qui consiste à estimer la diversité linguistique sans avoir connaissance des éco-régions ; nous dénommons cette approche évaluation du potentiel écolinguistique.

8.5.6.3 Potentiel écolinguistique

Cette méthode consiste à superposer les prédictions de niches et à observer les taux de prédictions additionnés. Afin d'optimiser la cohérence de l'approche, nous avons additionné séparément les ELN modélisées pour les groupes linguistiques papous et austronésiens. Pour ce faire, nous avons d'abord rendu binaire les prédictions en affectant la valeur 1 en tout point où la prédiction de présence était supérieure à 0.5 et une valeur nulle pour les autres points. De cette manière, nous ne conservons que les hautes probabilités de prédictions des diverses ELN. Puis nous avons additionné les couches de données ainsi modifiées, obtenant pour chaque pixel une valeur relative au nombre de niches prédites avec une valeur supérieure au seuil prédéfini.

Les résultats de ces superpositions sont présentés dans la figure 8.14. Les valeurs régionales des potentiels écolinguistiques coïncident relativement avec les valeurs de permissi-

Tableau 8.2 : Permissivités et potentiels écolinguistiques des motifs écologiques. La permissivité écologique est établie en fonction du nombre de niches ayant une expression géographique proche. Le potentiel écolinguistique est calculé en additionnant les niches écolinguistiques.

Motif écologique	Permissivité écologique	Potentiel écolinguistique
Hautes Terres	10	7-9
Centre	3	3-6
Sud-Est	3	3-6
Mangroves	5	2-4
Marécages	3	2-3
Unique [†]	1	1-3

(†) un motif unique correspond à la projection géographique d'une ELN de classe 2 ou 2*

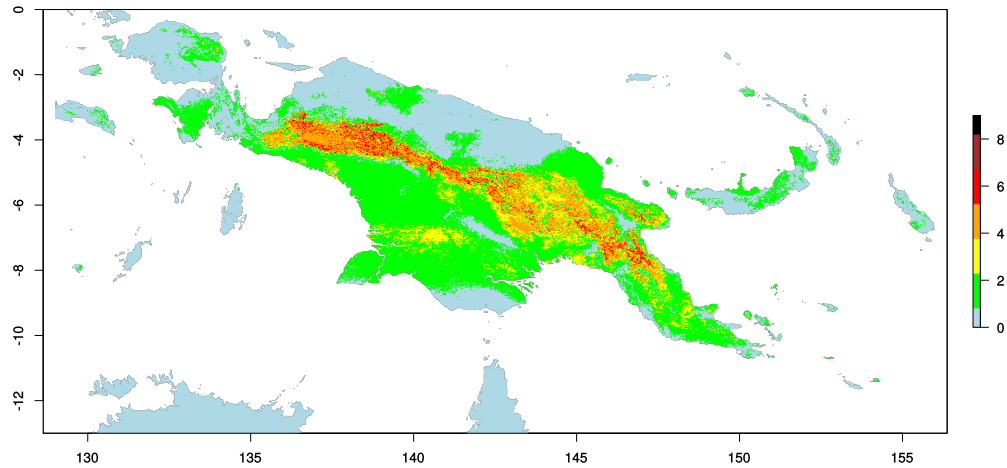
vité environnementale de chaque motif. Ainsi, les régions qui ont le potentiel écologique d'accueillir le plus grand nombre de langues papoues correspondent au motif des Hautes Terres avec un potentiel compris entre 7 et 9, suivi de la région qui équivaut aux motifs du Centre et du Sud-Est, dont les potentiels écolinguistiques de chacun varient de 3 à 6, puis par les Mangroves de 2 à 4 et les Marécages 2 à 3 (le tableau 8.2 permet de visualiser la ressemblance entre des résultats issus de chacune des méthodes).

Le tableau 8.2 nous laisse présager qu'il est possible de retrouver les motifs écologiques seulement à partir de l'analyse des ELN (observation de leur projection géographique et ACP). Ceci permet d'évaluer la diversité en fonction du risque écologique sans connaître les éco-régions ou par réciprocity évaluer le risque écologique encouru par des cultures en estimant le potentiel écolinguistique (inversement proportionnel).

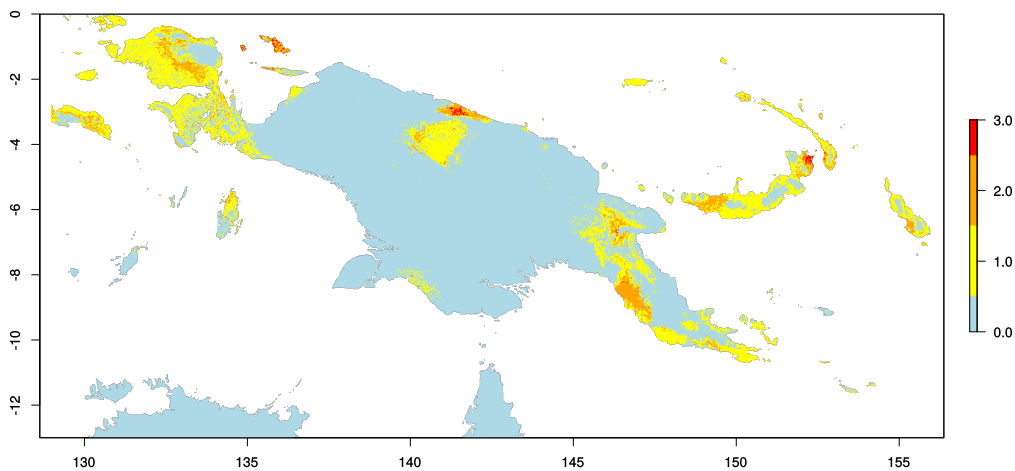
Lorsqu'on regarde en détail les potentiels écolinguistiques (figure 8.14), on observe pour les Austronésiens, qui ont en majorité des ELN uniques¹⁵, un maximum de régions avec un potentiel égal à 1, c'est à dire des régions dont l'environnement ne correspond qu'à une seule des niches écolinguistiques modélisées. Ce résultat est en accord avec l'attribution des classes préalablement effectuée. En revanche, certaines aires linguistiques dont l'ELN a été classée 2 ou 2*, donc qui ont par définition une permissivité écologique de 1, ont un potentiel écolinguistique égal à 2 ou 3, comme par exemple : les aires linguistiques de *Biak*, *New-Britain* et *Mekeo*. Ce résultat dévoile que des systèmes adaptatifs différents et présents en Nouvelle-Guinée peuvent convenir à un même environnement où le risque écologique contraint les populations à être unifiées et à restreindre la diversité linguistique. Nous suggérons que pour répondre à cette contrainte s'opère une sélection des langues pouvant être aléatoire ou non et pouvant varier en fonction de l'histoire des populations, des relations avec d'autres cultures, des catastrophes naturelles, etc. (voir section 1.6.6.4 pour comprendre comment peut se fixer un trait culturel).

Afin de vérifier l'hypothèse de Nettle [1998] sur la relation entre diversité linguistique et

15. la majorité des ELN des Austronésiens sont de classes 2 ou 2*, donc par définition non regroupables dans un motif écologique.



(a) Potentiel écolinguistique des groupes de langues papoues. Le nombre de groupes total est de 26 et le potentiel écolinguistique varie entre 0 et 9 ; sa valeur moyenne sur l'ensemble de la région est de 1.256.



(b) Potentiel écolinguistique des groupes de langues austronésiennes. Le nombre de groupes total est de 9 et le potentiel écolinguistique varie entre 0 et 3 ; sa valeur moyenne sur l'ensemble de la région est de 0.247.

FIGURE 8.14 : Potentiel écolinguistique de Nouvelle-Guinée. Ces cartes ont été établies en faisant la somme des prédictions d'ELN supérieures à 0.5 pour chaque groupe linguistique modélisées. Le potentiel écolinguistique est un indice indiquant le nombre d'ELN qui peuvent s'exprimer en un lieu donné.

risque écologique (section 8.1), dans la section suivante, nous allons comparer les potentiels écolinguistiques avec la durée de la saison de croissance des végétaux en Nouvelle-Guinée.

8.5.6.4 Corrélation entre richesse linguistique et saison de croissance des végétaux

Le risque écologique peut être défini comme étant l'inverse de la saison de croissance des végétaux [Nettle, 1996]. Cette mesure est nommée GS, pour *Growing Season*, et est exprimée en mois. Selon la formule de Le Houérou *et al.* [1989], un mois est inclus dans la GS si sa température moyenne est supérieure à 6°C et si le rapport précipitations sur températures est supérieur à 2. Si l'on applique cette formule pour la Nouvelle-Guinée, l'ensemble des lieux, hormis les régions au Sud présentant un climat Aw et correspondant à un biome dominé par la savane (figures 8.3(a) et 8.5), présente toute l'année les conditions favorables à la pousse des végétaux. Ces lieux ont donc un score de GS maximum (GS = 12). Nettle [1998] a calculé, pour la Papouasie Nouvelle-Guinée, une GS moyenne de 10.88, mais il faut considérer cette moyenne biaisée par la présence des prairies, savanes et terre arbustives tropicales (figure 8.5). Dans l'objectif d'apprécier les variations régionales, nous avons modifié la formule de la GS en multipliant les températures par 10. Nous avons nommé cette mesure GS₁₀ (voir fonction 8.1 et figure 8.15).

$$\text{SI} \left(\frac{\text{Précipitations}}{\text{Température} * 10} > 2 \text{ ET Température} > 6^{\circ}\text{C} \right) \quad (8.1)$$

ALORS{Le mois est inclus dans la GS}

Nous avons identifié, grâce à la projection géographique de la GS₁₀, les Hautes Terres comme étant la zone la plus prospère à la croissance des végétaux. C'est donc par conséquent la zone où le risque écologique est le moins élevé. Les Hautes Terres correspondent au motif précédemment décrit le plus représenté (10 ELN, dont celles des groupes : *Mek* 8.10(g), *Moni* 8.10(c) ou *Dani* 8.10(e)). Moins perceptible, les Monts Owen Stanley (Sud-Est de Papouasie) ont une GS₁₀ non nulle et correspondent également à des ELN de type 1 (*Binanderean* 8.10(s), *Southeast-Papuan* 8.10(t), *Angan* 8.10(u)).

Ainsi, après avoir estimé que les Hautes Terres possédaient les conditions environnementales permettant une haute diversité linguistique, nous montrons qu'elles possèdent également les conditions les plus favorables aux activités humaines. Ces résultats sont en accord avec l'hypothèse de Nettle [1998], énonçant que les environnements les plus cléments offrirent la possibilité aux populations de diverger linguistiquement. Cependant, des auteurs soutiennent que la corrélation entre PPN et diversité linguistique est attestée pour les sociétés de chasseurs-cueilleurs, mais qu'elle n'est pas valide pour les sociétés ayant une économie de production [Currie et Mace, 2012]. Or, les sociétés papoues ont un mode de subsistance basé sur l'horticulture et l'élevage. Au vu de cela, nous émettons l'hypothèse que la concordance entre la GS₁₀ (directement reliée à la PPN) et le motif des Hautes Terres (possédant la plus grande diversité linguistique de la région) pourrait être l'expression de la conservation d'un mode de subsistance ancien. La mise en évidence de

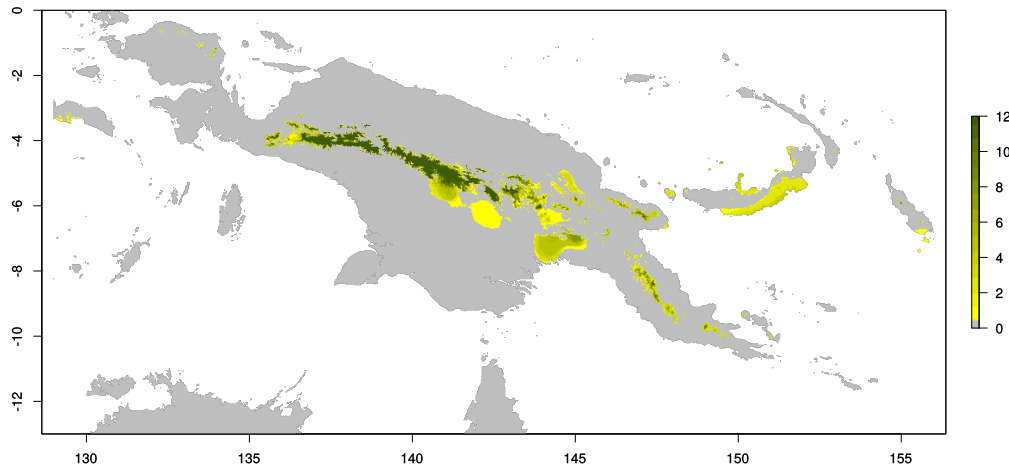


FIGURE 8.15 : Durée de croissance optimale des végétaux de Nouvelle-Guinée. Il s'agit de la GS₁₀, soit la saison de croissance des végétaux (GS : *Growing season*) de [Le Houérou et al. \[1989\]](#) où le rapport précipitations température a été affecté d'un facteur 10 pour apprécier les variations régionales.

cette conservation de niche suggère que l'installation des Papous dans les Hautes Terres remonterait au Pléistocène lors de la colonisation de l'île, puis que ces sociétés seraient restées très isolées depuis lors.

Le pendant de l'hypothèse de [Nettle](#) est que les environnements les plus hostiles sont moins permissifs à la dérive culturelle car les populations doivent être solidaires pour augmenter leurs chances de survie. Cela suggère aussi que les populations doivent agrandir leur territoire afin d'accéder au taux de ressource nécessaire à la survie du groupe et conformément à l'argument précédent, maintenir une intercompréhension dans tout ce territoire. Ce deuxième cas de figure correspond donc au cas 2 de l'ELNM où la niche prédite a une surface très semblable à celle de l'aire linguistique dont elle est la modélisation. Ainsi, lorsqu'on regarde les expressions des niches écolinguistiques des *Asmat-Kamoro* (09, figure 8.10(h)) ou des *Madang* (20, figure 8.10(q)), on s'aperçoit qu'elles sont situées dans des zones de GS₁₀ de valeur nulle, ce qui confirme également l'hypothèse sus-citée.

Nous constatons que l'hypothèse de [Nettle \[1996\]](#) ne se vérifie pas dans tous les cas de figure. Il est en effet des niches de type 1 qui ne coïncident pas avec la GS₁₀ comme celles correspondant aux motifs des Marécages [Marind (12) , Gogodala-Suki (28), Awyu-Domot (10)] ou à celui des Mangroves [Turama-Kikorian (26), Kiwai-Porome (27)].

Les étendues des niches écolinguistiques des Marécages et des Mangroves au-delà de leurs aires linguistiques ainsi que le chevauchement des ELN dans ces régions témoignent de leur potentiel à une grande diversité linguistique malgré un risque élevé si l'on considère

l'hypothèse de [Nettle \[1996\]](#). Ceci peut s'expliquer suivant deux hypothèses, qui vont dans le sens de celle énoncée par [Nettle \[1996\]](#) (c'est-à-dire que la diversité linguistique est corrélée au risque écologique), mais relevant de mécanismes plus complexes. La première hypothèse ferait intervenir des processus culturels alors que la seconde remet en cause l'évaluation du risque écologique établie précédemment (à savoir qu'il correspond à l'inverse de la saison de croissance des végétaux). Ces deux hypothèses, bien que l'une soit culturaliste et l'autre naturaliste, ne sont pas contradictoires et nous gardons à l'esprit que les deux peuvent être vraies conjointement ou indépendamment selon les cas.

L'hypothèse culturaliste suggérerait que la diversité linguistique soit plus élevée car les populations entretiendraient localement des échanges avec leurs voisins des régions à l'environnement plus favorable. Dans ce contexte, la cohésion sociale au sein de l'environnement pauvre serait délaissée et la compétition interne pour les ressources advenant accentuerait d'autant plus la diversité linguistique. Cette compétition et la relation avec le nombre de langues sont marquées par une agressivité accrue en certaines régions [[Schiefenhövel, 2001](#)]. Notons qu'une étude linguistique, portant sur la quantité d'emprunts lexicaux et syntaxiques aux langues voisines, apporterait des éléments permettant de valider cette hypothèse.

La deuxième hypothèse voudrait que les populations des régions correspondantes à ces motifs écologiques (qui ont des potentiels écolinguistiques variant de 2 à 4) ne soient pas gouvernées par les mêmes règles que celles qui vivent dans les Hautes Terres (c'est-à-dire que leur risque écologique est basé sur d'autres paramètres environnementaux). Les niches écolinguistiques concernées sont situées dans l'ACP en position p-eco4 et p-eco-5, il est donc possible de redéfinir les conditions qui leur sont favorables et d'en déduire leur risque écologique. l'ELNM permet donc, dans une certaine mesure, de mettre en évidence les changements de systèmes adaptatifs et d'en définir les nouvelles conditions environnementales favorables.

Dans la section suivante, nous verrons que la comparaison des ELN une à une, plutôt que groupées dans des motifs écologiques, confère une vision plus fine sur ce point.

8.5.6.5 *Overlap* des ELN

L'*overlap* (voir section 7.3) consiste à évaluer le pourcentage de prédictions communes entre deux ELN. Sa valeur varie entre 0 (deux niches entièrement distinctes) et 1 (deux niches parfaitement identiques). Nous avons utilisé le *package* R *phyloclim* [[Heibl et Calenge, 2013](#)] pour calculer tous les chevauchements entre les 35 groupes linguistiques (tableau 8.3). Nous avons ensuite transformé en distances les mesures obtenues ($\text{distance} = 1 - \text{overlap}$) et à partir de la matrice triangulaire de distance D de [Schoener](#), nous avons construit un arbre non raciné que nous nommons « phyloécolinguistique » (figure 8.16).

Nous constatons premièrement que les ELN des groupes linguistiques des deux familles de langues sont des groupes paraphylétiques. En effet, l'arbre de gauche (figure 8.16) montre que les groupes austronésiens se trouvent dans quatre clades. Nous relevons la présence de

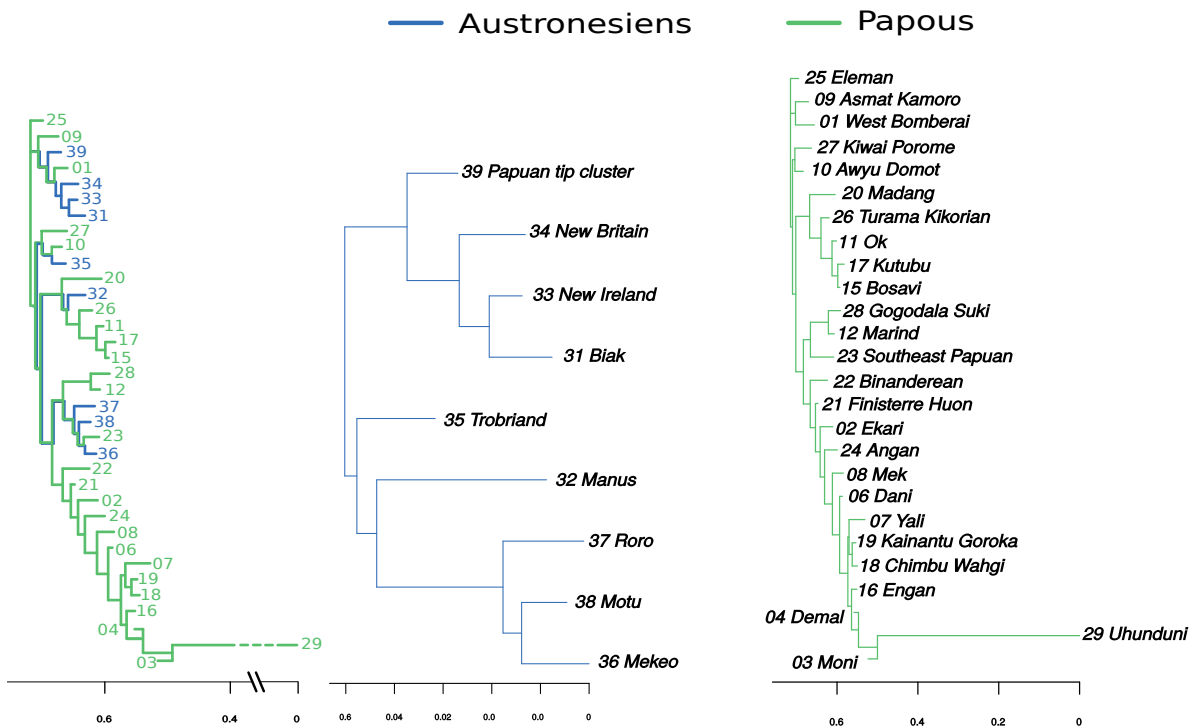


FIGURE 8.16 : Arbres phyloécologiques des groupes linguistiques de Nouvelle-Guinée. La famille papoue (en vert) et la famille austronésienne (en bleu) ne sont pas des groupes monophylétiques.

groupes papous dans ces clades et pouvons également noter que leur origine commune remonte au nœud le plus ancien. Il se distingue néanmoins un clade monophylétique chez les Papous composé des groupes linguistiques : 22, 21, 02, 24, 08, 06, 07, 19, 18, 16, 4, 3 et 29. Ce clade n'est pas sans rappeler les groupes linguistiques du motif des Hautes Terres (en position écologique p-eco1), en effet seuls les groupes 22 et 24 font partie d'un autre groupe (Sud-Est) et le groupe *Finisterre Huon* (21) dont la classification est délicate. Il est intéressant de constater l'éloignement considérable du groupe *Uhunduni* (29), linguistiquement classé comme un groupe isolé [Ross *et al.*, 2005]. Le motif Centre, composé des groupes 11, 15 et 17, constitue également un clade monophylétique. Par contre le groupe *Awyu-Domot* (10) est fortement éloigné de l'association Marind (12), Gogodala-Suki (28) que nous avons classés comme appartenant au motif des Marécages.

Tableau 8.3 : *Overlaps* entre ELN de Nouvelle-Guinée. Il s'agit de deux matrices triangulaires de similarité, la matrice supérieure est basée sur la distance D de Schoener [Schoener, 1968] alors que la matrice inférieure est basée sur la distance de Hellinger [Van der Vaart, 1998].

	29 Uhunduni	28 Gogodal Suki	27 Kiwai Porome	26 Turam Kikorian	25 Eleman	24 Angan	23 Southeast Papuan	22 Binderean	21 Finisterre Huon	20 Madang	19 Kaintu Goroka	18 Chimbu Wahgi	17 Kutubu	16 Engan	15 Bosavi	12 Marind	11 Ok	10 Awyu Domot	09 Asmat Kamoro	08 Mek	07 Yali	06 Dani	04 Demal	03 Moni	02 Ekari	01 West Bomberai	39 Papuan tip cluster	38 Motu	Roro	36 Mekeo 37	35 Trobriand	34 New Britain	33 New Ireland	32 Manus	31 Biak
31 Biak	-	0.65	0.78	0.71	0.6	0.56	0.55	0.62	0.64	0.7	0.58	0.42	0.5	0.57	0.53	0.59	0.63	0.57	0.52	0.52	0.49	0.49	0.47	0.49	0.5	0.57	0.63	0.57	0.52	0.52	0.49	0.49	0.47	0.49	0.9
32 Manus	0.9	-	0.61	0.65	0.63	0.47	0.49	0.5	0.55	0.6	0.56	0.44	0.49	0.55	0.44	0.61	0.66	0.67	0.68	0.49	0.63	0.48	0.59	0.45	0.45	0.63	0.58	0.5	0.44	0.53	0.66	0.73	0.6	0.43	0.05
33 New Ireland	0.96	0.88	-	0.75	0.61	0.54	0.55	0.61	0.71	0.79	0.59	0.42	0.48	0.56	0.53	0.58	0.68	0.57	0.51	0.51	0.47	0.48	0.45	0.48	0.49	0.56	0.69	0.56	0.51	0.53	0.62	0.52	0.54	0.51	0.03
34 New Britain	0.93	0.91	0.95	-	0.61	0.5	0.51	0.55	0.66	0.71	0.59	0.42	0.47	0.56	0.49	0.61	0.66	0.59	0.6	0.49	0.55	0.48	0.51	0.47	0.48	0.63	0.67	0.56	0.48	0.57	0.63	0.6	0.56	0.45	0.03
35 Trobriand	0.87	0.9	0.89	0.88	-	0.55	0.58	0.57	0.67	0.64	0.58	0.4	0.45	0.5	0.41	0.53	0.7	0.78	0.55	0.64	0.52	0.45	0.48	0.46	0.45	0.55	0.6	0.62	0.55	0.54	0.74	0.65	0.7	0.58	0.04
36 Mekeo	0.85	0.79	0.84	0.82	0.84	-	0.75	0.78	0.59	0.57	0.62	0.56	0.63	0.62	0.63	0.53	0.48	0.53	0.5	0.56	0.53	0.61	0.54	0.68	0.71	0.52	0.72	0.68	0.77	0.71	0.56	0.48	0.53	0.59	0.12
37 Roro	0.85	0.82	0.85	0.82	0.87	0.95	-	0.74	0.61	0.58	0.58	0.46	0.54	0.54	0.49	0.49	0.5	0.56	0.5	0.63	0.52	0.52	0.5	0.57	0.58	0.56	0.65	0.61	0.68	0.59	0.58	0.52	0.61	0.62	0.09
38 Motu	0.88	0.82	0.88	0.85	0.85	0.96	0.94	-	0.65	0.61	0.62	0.48	0.56	0.59	0.55	0.52	0.54	0.56	0.49	0.58	0.51	0.53	0.49	0.58	0.6	0.52	0.7	0.64	0.73	0.63	0.6	0.5	0.56	0.6	0.09
39 Papuan tip cluster	0.9	0.86	0.94	0.91	0.91	0.87	0.88	0.91	-	0.77	0.63	0.4	0.46	0.52	0.44	0.52	0.7	0.65	0.51	0.62	0.5	0.45	0.46	0.46	0.47	0.52	0.67	0.62	0.6	0.55	0.67	0.57	0.67	0.64	0.03
01 West Bomberai	0.93	0.88	0.97	0.93	0.9	0.85	0.86	0.87	0.96	-	0.65	0.43	0.49	0.57	0.5	0.58	0.69	0.62	0.54	0.56	0.52	0.49	0.49	0.49	0.5	0.55	0.71	0.6	0.54	0.57	0.66	0.57	0.59	0.54	0.04
02 Ekari	0.85	0.84	0.86	0.86	0.86	0.89	0.86	0.88	0.88	-	0.66	0.7	0.72	0.57	0.67	0.56	0.57	0.61	0.52	0.61	0.64	0.62	0.6	0.62	0.54	0.74	0.62	0.61	0.67	0.62	0.57	0.55	0.5	0.19	
03 Moni	0.75	0.76	0.74	0.74	0.75	0.85	0.78	0.79	0.73	0.75	0.92	-	0.79	0.72	0.66	0.67	0.37	0.38	0.56	0.34	0.58	0.75	0.63	0.66	0.68	0.46	0.59	0.52	0.51	0.62	0.42	0.47	0.37	0.34	0.28
04 Demal	0.81	0.8	0.8	0.79	0.79	0.89	0.82	0.85	0.78	0.81	0.93	0.96	-	0.82	0.7	0.71	0.43	0.43	0.6	0.41	0.6	0.77	0.65	0.69	0.7	0.51	0.67	0.6	0.57	0.72	0.49	0.5	0.42	0.4	0.26
06 Dani	0.85	0.83	0.85	0.84	0.82	0.89	0.83	0.86	0.82	0.85	0.94	0.94	0.98	-	0.72	0.74	0.5	0.49	0.6	0.44	0.59	0.77	0.62	0.68	0.7	0.55	0.73	0.61	0.56	0.74	0.54	0.52	0.45	0.42	0.2
07 Yali	0.84	0.75	0.83	0.79	0.74	0.89	0.8	0.84	0.77	0.8	0.87	0.9	0.93	0.94	-	0.62	0.42	0.37	0.48	0.36	0.47	0.71	0.52	0.71	0.73	0.49	0.67	0.57	0.54	0.67	0.43	0.39	0.35	0.35	0.21
08 Mek	0.86	0.88	0.86	0.87	0.83	0.84	0.8	0.82	0.81	0.86	0.92	0.91	0.94	0.95	0.89	-	0.53	0.51	0.7	0.42	0.65	0.7	0.68	0.6	0.59	0.59	0.68	0.55	0.5	0.63	0.54	0.58	0.47	0.39	0.18
09 Asmat Kamoro	0.89	0.91	0.91	0.91	0.93	0.79	0.81	0.83	0.93	0.92	0.85	0.72	0.76	0.81	0.73	0.83	-	0.7	0.51	0.54	0.48	0.43	0.44	0.43	0.43	0.5	0.6	0.57	0.48	0.54	0.75	0.59	0.61	0.5	0.03
10 Awyu Domot	0.85	0.91	0.86	0.88	0.96	0.82	0.85	0.83	0.91	0.89	0.85	0.73	0.77	0.8	0.7	0.82	0.93	-	0.57	0.68	0.54	0.43	0.5	0.45	0.44	0.55	0.59	0.62	0.55	0.52	0.76	0.69	0.71	0.59	0.04
11 Ok	0.82	0.92	0.82	0.87	0.85	0.82	0.8	0.81	0.82	0.84	0.89	0.86	0.87	0.88	0.79	0.93	0.82	0.86	-	0.46	0.81	0.59	0.77	0.53	0.52	0.63	0.6	0.54	0.48	0.58	0.57	0.72	0.55	0.4	0.12
12 Marind	0.82	0.81	0.83	0.8	0.9	0.86	0.9	0.87	0.89	0.85	0.82	0.68	0.74	0.76	0.7	0.74	0.84	0.92	0.76	-	0.45	0.4	0.41	0.42	0.42	0.47	0.55	0.61	0.61	0.47	0.62	0.51	0.65	0.76	0.04
15 Bosavi	0.79	0.9	0.79	0.84	0.83	0.83	0.82	0.81	0.81	0.82	0.88	0.86	0.87	0.87	0.78	0.91	0.8	0.84	0.97	0.75	-	0.6	0.84	0.54	0.54	0.62	0.59	0.54	0.49	0.59	0.55	0.73	0.55	0.41	0.13
16 Engan	0.8	0.8	0.79	0.79	0.78	0.87	0.79	0.81	0.76	0.79	0.9	0.95	0.96	0.96	0.93	0.93	0.76	0.76	0.86	0.7	0.87	-	0.65	0.8	0.78	0.52	0.64	0.59	0.55	0.69	0.47	0.49	0.41	0.37	0.23
17 Kutubu	0.78	0.87	0.78	0.82	0.8	0.84	0.81	0.81	0.79	0.8	0.88	0.89	0.9	0.89	0.81	0.92	0.78	0.81	0.96	0.72	0.98	0.9	-	0.58	0.57	0.6	0.6	0.53	0.5	0.61	0.51	0.67	0.5	0.38	0.16
18 Chimbu Wahgi	0.8	0.77	0.79	0.78	0.78	0.91	0.84	0.85	0.78	0.79	0.87	0.91	0.92	0.91	0.94	0.88	0.74	0.76	0.82	0.75	0.83	0.97	0.86	-	0.84	0.52	0.63	0.61	0.61	0.67	0.48	0.48	0.44	0.42	0.23
19 Kaintu Goroka	0.81	0.76	0.79	0.79	0.78	0.93	0.85	0.87	0.79	0.8	0.88	0.92	0.93	0.93	0.95	0.88	0.74	0.75	0.82	0.75	0.83	0.96	0.86	0.98	-	0.51	0.67	0.63	0.63	0.71	0.48	0.46	0.42	0.41	0.21
20 Madang	0.86	0.89	0.86	0.89	0.84	0.83	0.85	0.83	0.83	0.85	0.83	0.76	0.8	0.83	0.8	0.86	0.82	0.85	0.89	0.78	0.88	0.82	0.86	0.82	0.81	-	0.61	0.55	0.47	0.55	0.56	0.62	0.55	0.41	0.07
21 Finisterre Huon	0.91	0.87	0.92	0.92	0.88	0.94	0.91	0.93	0.92	0.92	0.94	0.88	0.92	0.94	0.92	0.92	0.88	0.87	0.89	0.85	0.88	0.91	0.88	0.9	0.92	0.89	-	0.72	0.68	0.74	0.64	0.56	0.56	0.53	0.1
22 Binderean	0.85	0.82	0.85	0.85	0.89	0.92	0.88	0.89	0.89	0.87	0.88	0.83	0.87	0.88	0.85	0.85	0.86	0.88	0.83	0.87	0.84	0.87	0.83	0.89	0.9	0.86	0.94	-	0.71	0.7	0.64	0.54	0.57	0.54	0.1
23 Southeast Papuan	0.83	0.78	0.83	0.81	0.84	0.96	0.91	0.94	0.88	0.84	0.89	0.82	0.86	0.85	0.84	0.82	0.8	0.84	0.8	0.89	0.81	0.83	0.81	0.88	0.9	0.8	0.93	0.93	-	0.65	0.56	0.48	0.54	0.63	0.11
24 Angan	0.84	0.83	0.84	0.85	0.84	0.93	0.86	0.89	0.85	0.85	0.91	0.89	0.93	0.94	0.91	0.89	0.83	0.83	0.86	0.79	0.87	0.92	0.89	0.91	0.93	0.84	0.95	0.93	0.9	-	0.58	0.54	0.49	0.45	0.14
25 Eleman	0.88	0.91	0.89	0.9	0.95	0.84	0.86	0.87	0.92	0.9	0.88	0.77	0.81	0.85	0.76	0.84	0.95	0.95	0.86	0.89	0.84	0.8	0.82	0.79	0.79	0.85	0.9	0.9	0.85	0.87	-	0.67	0.7	0.56	0.05
26 Turam Kikorian	0.82	0.94	0.83	0.88	0.9	0.81	0.83	0.82	0.85	0.85	0.86	0.79	0.81	0.83	0.72	0.87	0.87	0.92	0.94	0.82	0.94	0.8	0.92	0.78	0.77	0.89	0.86	0.84	0.8	0.84	0.91	-	0.69	0.46	0.09
27 Kiwai Porome	0.84	0.88	0.85	0.86	0.93	0.82	0.87	0.85	0.91	0.88	0.84	0.71	0.75	0.78	0.69	0.79	0.89	0.94	0.85	0.91	0.85	0.73	0.81	0.75	0.74	0.85	0.86	0.85	0.83	0.8	0.93	0.92	-	0.64	0.06
28 Gogodal Suki	0.81	0.77	0.83	0.78	0.87	0.86	0.89	0.88	0.9	0.84	0.82	0.67	0.74	0.75	0.7	0.72	0.82	0.87	0.73	0.96	0.72	0.68	0.7	0.73	0.74	0.74	0.85	0.84	0.89	0.77	0.85	0.77	0.9	-	0.04
29 Uhunduni	0.21	0.21	0.18	0.17	0.2	0.33	0.26	0.27	0.17	0.18	0.41	0.52	0.5	0.43	0.44	0.4</																			

L'analyse de cet arbre met en exergue des niches relativement proches. Cette proximité dénote le caractère plésiomorphe de certains systèmes adaptatifs. Effectivement, si nous considérons les Austronésiens comme une société tournée vers la mer ayant hérité de nombreuses pratiques et technologies du continent asiatique et ayant subi l'influence des populations croisées dans la vaste zone sur laquelle ils se sont étendus, leur système adaptatif s'est trouvé alors plus souvent « réactualisé », il est donc pour ainsi dire plus « jeune ». À l'inverse, considérant les Papous (notamment ceux des Hautes Terres) comme des populations moins mobiles, isolées et implantées depuis une très longue période, nous pouvons estimer que les évolutions de leurs systèmes adaptatifs sont majoritairement façonnées par des phénomènes de dérives aléatoires. L'arrivée relativement récente des Austronésiens (il y a environ 7000 ans, section 8.3.2) a initié des modifications au niveau linguistique, économique et culturel (par exemple [Merriwether *et al.*, 1999; Terrill, 2003]). Cette acculturation débute relativement récemment par rapport aux premiers peuplements de la Nouvelle-Guinée. Ceci laisse présager que les populations et leurs systèmes adaptatifs ont pu se mettre en place avant cet événement et que des phénomènes de dérives culturelles (dont la dérive linguistique) « tolérés » par l'environnement ont pu avoir lieu. Suivant ces hypothèses les positions relatives des groupes papous par rapport aux groupes austronésiens dans l'arbre phyloécolinguistique nous permettent d'estimer leur proximité écologique et donc le degré d'acculturation possible. L'analyse des arbres séparés des deux familles apporte une information sur leur dérive culturelle interne et donc une estimation sur leur date d'implantation.

Les caractères d'une langue évoluant selon plusieurs vitesses, dans le futur il serait intéressant de reconstruire des arbres phylolinguistiques basés sur des caractères à mutation lente et de les comparer avec les arbres phyloécolinguistiques. Il serait également intéressant de comparer ces résultats avec des arbres phylogénétiques et des *clusters* établis selon d'autres marqueurs culturels comme des éléments matériels qui nous permettraient d'intégrer des vestiges archéologiques. Ces études supplémentaires nous permettraient d'intégrer des dimensions temporelles, culturelles et biologiques à notre étude.

Pour étayer nos propos quant à la différence de systèmes adaptatifs entre Papous et Austronésiens, nous avons regroupé les 35 ELN par familles linguistiques et avons étudié leurs positions écologiques respectives.

8.5.7 Différences Papous-Austronésiens

La majorité des ELN papoues sont de catégorie 1, huit ELN austronésiennes sur neuf sont dans le cas 2 dont cinq sont dans le cas 2*. Le potentiel écolinguistique moyen¹⁶ papou est de 1.256 et celui des Austronésiens est de 0.247. À première vue, ces résultats suggèrent qu'il y a une différence entre ces deux familles linguistiques au niveau de l'exploitation des ressources, ainsi que des milieux occupés : comparés aux Papous, les Austronésiens

16. moyenne du potentiel linguistique sur toute la région étudiée

seraient confrontés à un risque écologique supérieur. Toutefois nous devons considérer la possibilité que certaines niches austronésiennes classées 2* puissent être classées comme des niches de catégories 1 du fait de la forte mobilité des Austronésiens que leur confèrent leurs excellentes capacités de navigation.

Dans la figure 8.17, nous avons rassemblé les groupes linguistiques selon leur famille afin de mettre en évidence la position écologique de ces dernières. Nous observons que les familles papoue et austronésienne n'occupent pas la même position écologique. La forme oblongue de l'ellipse d'inertie des Austronésiens est due à l'alignement des groupes linguistiques. Cet alignement est parallèle au deuxième axe de l'ACP qui reflète l'intensité et la saisonnalité des précipitations selon un gradient environnemental (voir figure 8.12). La

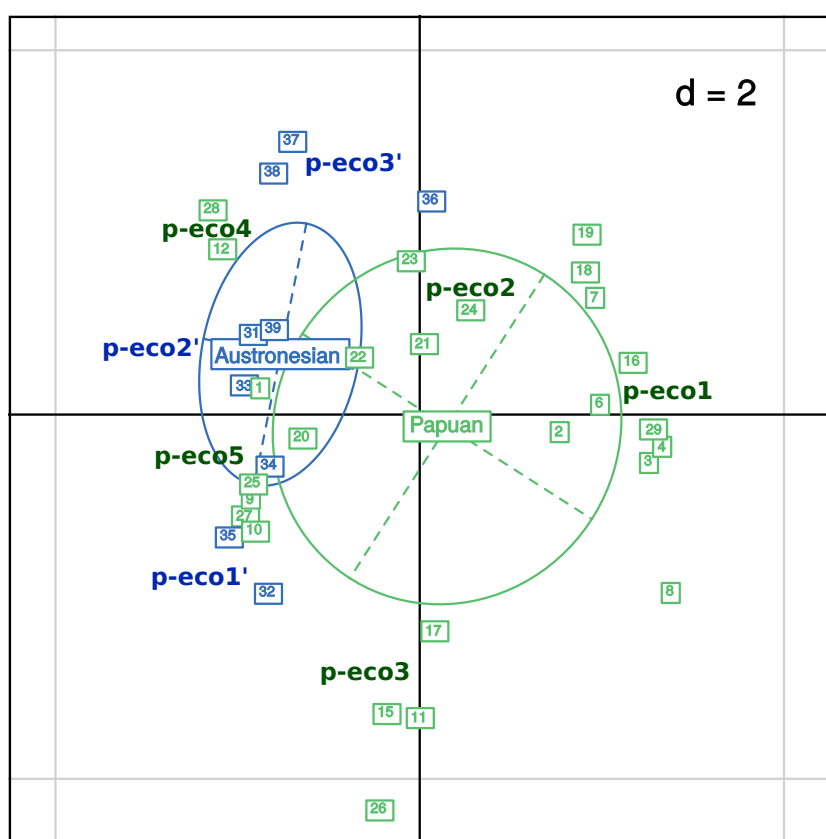


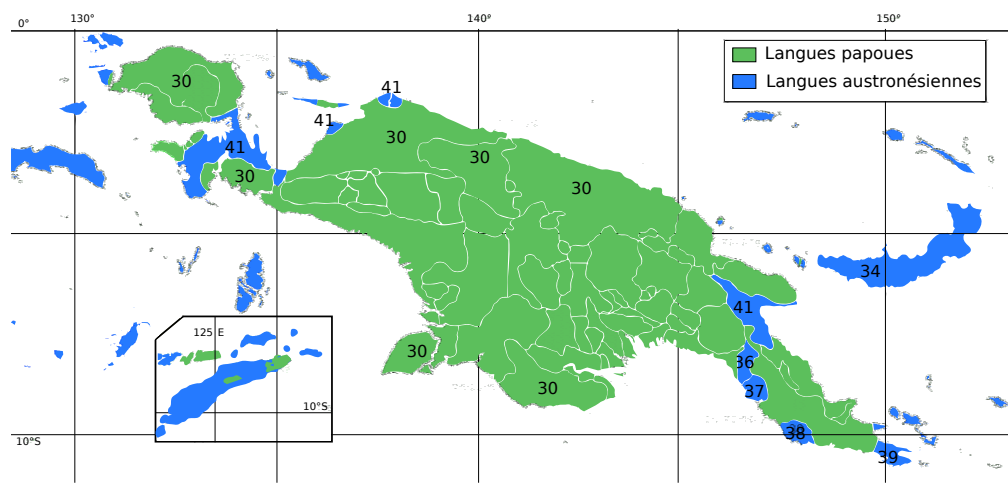
FIGURE 8.17 : Comparaison des positions écologiques occupées par les Papous et par les Austronésiens. En vert les informations relatives aux Papous, en bleu celles relatives aux Austronésiens. Les numéros correspondent aux groupes linguistiques. En gras sont indiquées les positions écologiques (p-eco1 à p-eco5 pour les Papous, p-eco1' à p-eco3' pour les Austronésiens).

forme allongée de l'ellipse des Austronésiens suggère aussi que ces groupes linguistiques ont une position écologique similaire vis-à-vis des températures. L'inertie des Papous est quant à elle en position centrale. La forme quasi circulaire de l'ellipse qui la représente

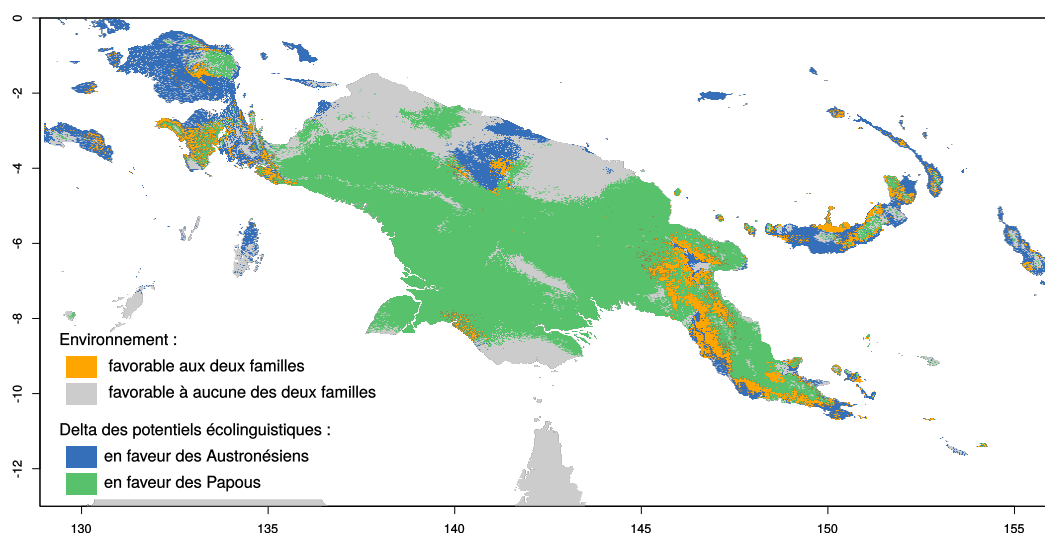
est due à la diversité des positions écologiques des groupes linguistiques qui composent cette famille (voir la description des p-eco section 8.5.6.1). La constitution d'ensembles de groupes linguistiques papous selon leur position écologique permettrait de proposer une formule spécifique à chacun. Il serait judicieux pour cette tâche de s'entourer d'agronomes et pédologues.

L'ACP dévoile donc une diversité d'ELN supérieure chez les Papous reflétant leur potentiel à exploiter un plus grand nombre de milieux. Cette différence peut être expliquée par l'acquisition de systèmes adaptatifs spécifiques à certains milieux chez les Papous qui auraient pu se développer avec l'accès et la prise de possession de ces terres antérieurement à l'arrivée des Austronésiens. L'observation d'ELN de groupes papous disposées le long du gradient de précipitations laisse penser que ces derniers occupent une position écologique proche de celle des Austronésiens et étaye la thèse du peuplement plus tardif de ces derniers, les terres les moins favorables étant occupées en dernier [Coddington et Jones, 2013].

Le chevauchement réduit des ellipses des deux familles montre que seulement une faible partie des groupes de chacune peuvent se trouver dans cette situation et être donc en compétition pour un territoire seyant à l'écologie de chacune. Le fait d'occuper deux niches différentes est une solution pour éviter la compétition et permettre la coexistence (voir chapitre 1). Dans la figure 8.18(b), on distingue les régions les plus favorables aux Papous de celles les plus favorables aux Austronésiens. Pour obtenir cette figure, nous avons normalisé chacune des cartes de potentiels écolinguistiques (figure 8.14) et établi leurs différences. Nous dénommons la carte des différences « delta des potentiels écolinguistiques », il permet d'évaluer si l'environnement est potentiellement plus favorable à une famille qu'à une autre. Nous trouvons de grandes similitudes entre cette carte et celle dessinée par les linguistes (voir [Schiefenhövel, 2014]). En effet, plusieurs aires linguistiques non incluses dans l'analyse ont été prédites comme ayant un environnement potentiellement favorable à accueillir des populations parlant une langue de la famille linguistique identifiée pour ces régions. Le fait que les aires linguistiques, identifiées comme *other Papuan* (30) et *other Austronesian* (41), soient, dans la plupart des cas, correctement attribuées par l'ELNM, nous permet de valider notre approche. Sur la figure 8.18(b), il est possible d'observer géographiquement à la fois les régions les plus propices à chacune des deux familles (colorées en bleu ou en vert), les lieux possédant un environnement approprié aux deux familles (en orange) et ceux qui ne conviennent à aucun des groupes modélisés (en gris). Cette figure nous permet d'apprécier le chevauchement des niches qui peut refléter les zones les plus convoitées et dans lesquelles des conflits territoriaux (pour des ressources environnementales) pourraient exister. À l'opposé, les régions dont le delta du potentiel écolinguistique est prédit comme favorable aux Austronésiens sont, à l'exception d'une région au Centre Nord, occupées par des locuteurs austronésiens. Nous défendons l'hypothèse d'un comportement opportuniste de ces derniers, car arrivant dans une terre totalement occupée par des papous il y a environ 7000 ans (section 8.3), il devait être alors préférable de s'installer dans des régions dont l'environnement était le moins propice aux populations déjà en place, car elles étaient certainement défendues avec moins d'ardeur que les autres.



(a) Distribution des langues papoues et austronésiennes de Nouvelle-Guinée (d'après [Schiefenhövel \[2014\]](#), carte modifiée). *Other Papuan*(30) *Mekeo* (36), *Roro* (37), *Motu* (38), *Papuan TIP cluster* (39) et *Other Austronesian* (41).



(b) Delta de potentiels écolinguistiques entre langues papoues et austronésiennes. Le delta a été calculée en faisant la différence entre les potentiels écolinguistiques papous et les potentiels écolinguistiques des austronésiens.

FIGURE 8.18 : Comparaison de la répartition des langues papoues et austronésiennes avec le delta des potentiels écolinguistiques. On observe une grande ressemblance entre la carte de populations connues et celle obtenues à partir des ELN. Nous devons noter que les aires linguistiques *other Austronesian* (41) et *other Papuan* (30) n'ont pas été incorporées dans l'ELNM et sont dans la majorité des cas prédites comme ayant un environnement favorable à leur famille linguistique respective.

Cet opportunisme peut expliquer la présence d'Austronésiens dans l'île principale qui est majoritairement peuplée de Papous.

8.5.7.1 Des Austronésiens en terres papoues

Plusieurs populations austronésiennes se trouvent sur l'île principale de Nouvelle-Guinée comme les *Roro* (37), *Mekeo* (36), *Motu* (38), *Papuan-tip-clusters* (39) et certains *other Austronesian* (41, « dans le cou de l'oiseau » et à l'Est de l'île). Les quatre ELN des Austronésiens de l'Est de l'île que nous avons modélisées (ELN des groupes 36, 37, 38 et 39) ont chacune une région fortement prédite en commun se trouvant au Sud de la péninsule Huon. Les distances entre cette région et les aires linguistiques respectives et *Motu* expliquent leur classement en 2*. Le groupe linguistique *Mekeo* présente quant à lui, selon les critères que nous avons défini, une ELN classée 1. Cette région de prédictions communes est frontalière avec l'aire linguistique *Mekeo* ce qui confère à l'ELN un classement dans la catégorie 1 au lieu d'être dans la catégorie 2*. Si le classement en 2 renseigne sur l'exploitation du milieu ou sur le risque écologique, il est fort probable que les trois groupes bien que « spécialisés » aient des systèmes adaptatifs relativement similaires. Il est intéressant de constater que cette région est occupée dans sa partie sud par un groupe d'Austronésiens (que nous n'avons pas défini, et dont nous n'avons pas reconstruit la niche, *Other Austronesian* (41)), lesquels empêchent possiblement l'extension des Mekeos dans cette zone. La partie septentrionale de cette zone, commune aux ELN, est quant à elle occupée par les *Finisterre-Huon* dont l'ELN a une projection géographique diffuse dans une très grande partie de Nouvelle-Guinée. La présence antérieure de ces derniers a pu jouer également dans la non-annexion de cette zone par des Austronésiens.

Nous remarquons en effet que les groupes linguistiques papous en position écologique p-eco5 (1, 9, 10, 20, 25 et 27) s'intercalent entre les groupes linguistiques des austronésiens en positions p-eco1' (33, 34, 35) et p-eco2' (31, 33, 39). Cela laisse présager que si l'espace n'était pas occupé par des Papous, il pourrait l'être par des Austronésiens. Ce constat soulève de nouveaux questionnements sur la présence des Papous dans cette région. Leur présence dans ces lieux est-elle antérieure à l'arrivée des Austronésiens ou ont-ils supplanté des populations austronésiennes ? Nous pouvons également nous demander si l'adoption de nouveaux comportements des groupes papous (induits par la présence ou la proximité des Austronésiens) leur aurait permis de se maintenir en limite de leur niche écologique, s'il s'agit d'une convergence des systèmes adaptatifs ou si l'on assiste à un décalage de niche ayant lieu afin d'éviter la compétition (voir chapitre 1).

8.6 Perspectives

Des projections (voir section 2.2.2) des ELN obtenues pour la zone d'étude précédemment décrite (section 8.5.3) dans d'autres régions et à d'autres époques pourraient être

très instructives afin d'expliquer la rémanence actuelle de langues papoues de régions éloignées. Elles permettraient par exemple d'identifier s'il existe une explication écologique à la présence de langues papoues des îles de *Timor* et d'*Alor*. Les projections des différentes ELN des Austronésiens sur le continent sud-américain apporteraient un éclairage quant aux lieux possédant un environnement convenable à ces cultures dans ces régions. Une telle étude permettrait d'apporter des éléments supplémentaires à l'écart ou l'adoption de théories décrivant un possible peuplement de ce continent par les Austronésiens.

Dans le contexte d'une étude ethnoarchéologique plus large, nous souhaitons coupler l'ELNM à une ECNM basée sur des marqueurs culturels matériels afin de créer un référentiel qui permettrait à terme d'appréhender l'ethnicité de populations disparues. Nous projetons également de confronter ces résultats à ceux issus d'autres disciplines comme la génétique et la linguistique afin d'identifier les diffusions démiques, linguistiques et culturelles.

Enfin, la mise en évidence des deux catégories de niches (section 8.5.6.2) permet une discussion sur les frontières culturelles. On assiste à un morcellement culturel dans le premier cas et à une unification culturelle dans le second. Une étude sur la diffusion des multiples variations culturelles dans les deux situations nous renseignerait sur la perméabilité et la mobilité de ces frontières.

Synthèse du chapitre : Impact de l'environnement sur la diversité linguistique de Nouvelle-Guinée

La comparaison des surfaces entre les aires linguistiques et leur ELN nous permet d'apprécier la densité culturelle et par conséquent, selon l'hypothèse de [Nettle \[1996\]](#), d'estimer le risque écologique. Les ELN des groupes linguistiques des Hautes Terres sont très similaires et constituent ce que nous avons nommé un motif écologique. Ce dernier présente une forte similitude avec la production primaire nette (PPN) des écosystèmes. Cette ressemblance confirme l'hypothèse précédemment évoquée qui stipule que la diversité culturelle est plus importante dans les environnements favorables, c'est-à-dire en rapport avec la durée de la saison de croissance des végétaux.

Après avoir mis en évidence la présence d'autres motifs écologiques ne reflétant pas la PPN, nous suggérons que la formule du risque écologique n'est pas universelle. Nous indiquons alors que l'observation des positions écologiques des niches écolinguistiques peut aider à dresser ces formules.

Dans ce chapitre, une distinction entre les ELN papoues et austronésiennes a été mise en exergue. Nous émettons des hypothèses culturelles et environnementales pour expliquer cette différence.

Concluant sur le fait que la complexité de la géographie culturelle d'une région est liée à l'histoire des populations ainsi qu'aux adaptations à l'environnement, nous envisageons d'allier cette étude à d'autres portant sur l'archéologie matérielle et la génétique afin d'apporter un nouveau regard sur l'histoire du peuplement de Nouvelle-Guinée.

Impact du changement climatique sur les populations scandinaves du Groenland

Préambule

Pendant l'optimum climatique médiéval, les Vikings bénéficièrent d'un climat favorable au fondement de colonies dans les régions subarctiques d'Islande, du Groenland et du Labrador. Dans ces régions, la culture viking connut une période faste, autour du XI^e siècle, puis s'est progressivement dégradée et finit par disparaître près de quatre siècles plus tard. Il existe plusieurs hypothèses politiques, économiques, sanitaires ou écologiques pour expliquer l'effondrement de cette société, mais la plupart des auteurs s'accorde à dire que la détérioration climatique due à l'avènement du petit âge glaciaire serait une des causes majeures de l'abandon de ces régions. À l'aide des algorithmes prédictifs combinant des données archéologiques et paléoenvironnementales, nous sommes capables de reconstruire la niche écoculturelle de ces Vikings et d'apprécier son évolution au cours du temps. Ainsi, nous pouvons déterminer combien la détérioration climatique durant le petit âge glaciaire fut responsable de la chute de cette civilisation dans ces régions.

Sommaire

9.1	La détérioration climatique du PAG	188
9.2	L'histoire des Vikings des régions subarctiques	189
9.3	ECNM des Vikings des régions subarctiques	196
9.4	Interactions culturelles	223

9.1 La détérioration climatique du PAG

L'Optimum Médiéval Climatique (OCM¹), est suivi d'une période de détérioration climatique appelée Petit Âge Glaciaire (PAG²) [Mann, 2002].

Ces variations climatiques sont attestées par des *proxies*³ tels que des cernes d'arbres, des carottes de glace, des anneaux de coraux, des sédiments lacustres ou encore des calcites de spéléothèmes. D'autre part il a été recensé de nombreux témoignages ou représentations historiques (peintures, dates de récoltes, date des vendanges, etc) [Jones et Mann, 2004; Yiou et Masson-Delmotte, 2005] évoquant le ressenti d'une détérioration climatique à partir de l'an 1300.

Les différences entre PAG et OCM sont bien marquées dans les reconstructions climatiques avec des *extrema* autour des x^e et xi^e siècles pour la période tempérée et entre les xvii^e et xviii^e siècles pour la période froide [Servonnat, 2010]. Mann (figure 9.1), montre que ces phases climatiques n'ont pas eu le même impact sur tout l'hémisphère Nord. Cependant l'hétérogénéité du phénomène ne peut être vérifié du fait du manque de *proxy* [Jones et al., 2009].

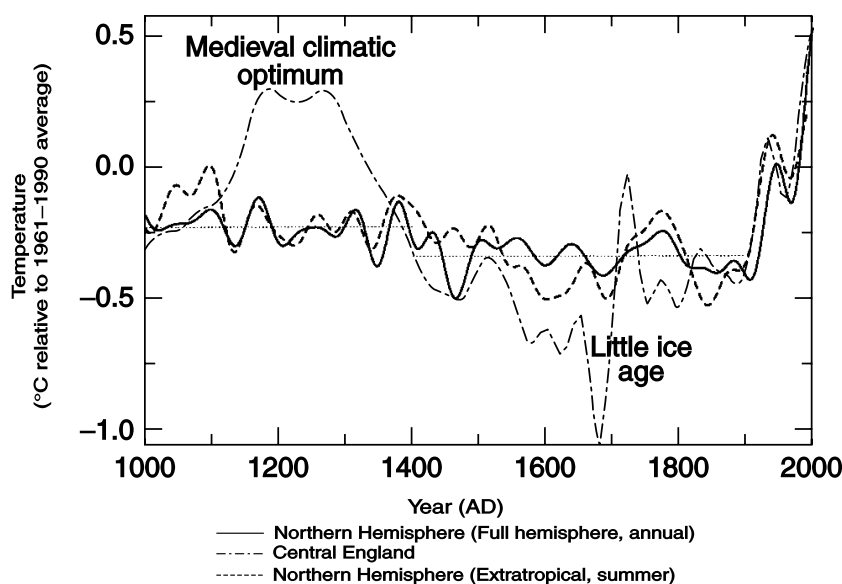


FIGURE 9.1 : Anomalies des températures du dernier millénaire, d'après Mann [2002].

Les Vikings ont évolué durant ces deux périodes. Après avoir relaté leur histoire dans les régions subarctiques, nous essaierons d'apprécier l'impact de cette détérioration climatique sur leur civilisation.

1. en anglais, MCO : *Medieval Climatic Optimum*, période chaude, comprise entre 900 et 1300 AD
2. en anglais, LIA : *Little Ice Age*, période froide
3. Un *proxy* est une mesure en un lieu donné permettant d'évaluer un paramètre.

9.2 L'histoire des Vikings des régions subarctiques

Dans ce document nous employons le terme de « Viking », bien qu'il signifie commerçant pillard à l'origine, pour désigner l'ensemble de la population scandinave ayant vécu entre le IX^e et le XVI^e siècle. Nous verrons que malgré l'image quelque peu faussée des Scandinaves qui nous est parvenue, leur histoire est quant à elle bien documentée à la fois par des documents écrits, comme les sagas, mais aussi par des vestiges archéologiques. Ainsi les connaissances que nous avons sur leurs mœurs et leur histoire vont nous permettre une bonne interprétation des résultats issus de l'ECNM.

9.2.1 Les Vikings, une épopée nordique

L'Âge viking s'étala sur plusieurs siècles. D'origine norvégienne, leur emprise géographique fut une des plus vastes. Redoutables conquérants et pillers mais aussi agriculteurs, fermiers et commerçants, ils atteignirent les frontières de la Russie d'une part et l'Amérique du Nord d'autre part. Certains faits historiques recensent également des raids en Méditerranée ou en Russie [Béraud-Villars, 1951; Maillefer, 2005]. Ayant fait régner la terreur sur l'Europe entière, leur faits d'armes sont reportés dans des récits médiévaux qui leur dépeignent une image d'effroyables guerriers, de nos jours encore gravée dans notre imaginaire.

9.2.2 Redoutables guerriers, mythe ou réalité ?

Notre vision des Vikings est souvent faussée par leur réputation de guerriers sanguinaires suite aux nombreux traumatismes infligés après avoir mis à sac des grandes villes, des régions et des pays entiers comme Nantes, la Normandie ou encore l'Angleterre [Winroth, 2014]. Cette réputation devait émaner pour grande part des Vikings eux-mêmes, leur stratégie militaire étaient alors d'instaurer la terreur en ornant leurs vaisseaux de guerre de têtes de dragons (ce qui leur valu plus tard l'appellation de drakkar). Également dans le même objectif, les têtes coupées et les corps empalés étaient de mise dans les représentations iconographiques (figure 9.2) et dans l'oralité (saga, récits épiques, poèmes et chansons).

Cependant, les épisodes guerriers étaient brefs. La plupart du temps, ces populations se consacraient aux travaux de la ferme ou aux échanges commerciaux, principalement maritimes [Buckland *et al.*, 1996]. Les sagas nordiques témoignent d'ailleurs de ces faits en décrivant leurs héros comme des hommes partis en *terra incognita* et ayant fait fortune suite à leur dur labeur [Gravier, 1955; Sveinsson *et al.*, 1935].

Nos travaux consistent à établir les relations Homme environnement, donc nous focaliserons notre intérêt sur les lieux de vie de ces populations et non sur leur aire d'influence. Pour ce faire nous nous basons sur des écrits et des vestiges archéologiques qui témoignent de leurs modes de subsistance et de leurs lieux de vie.

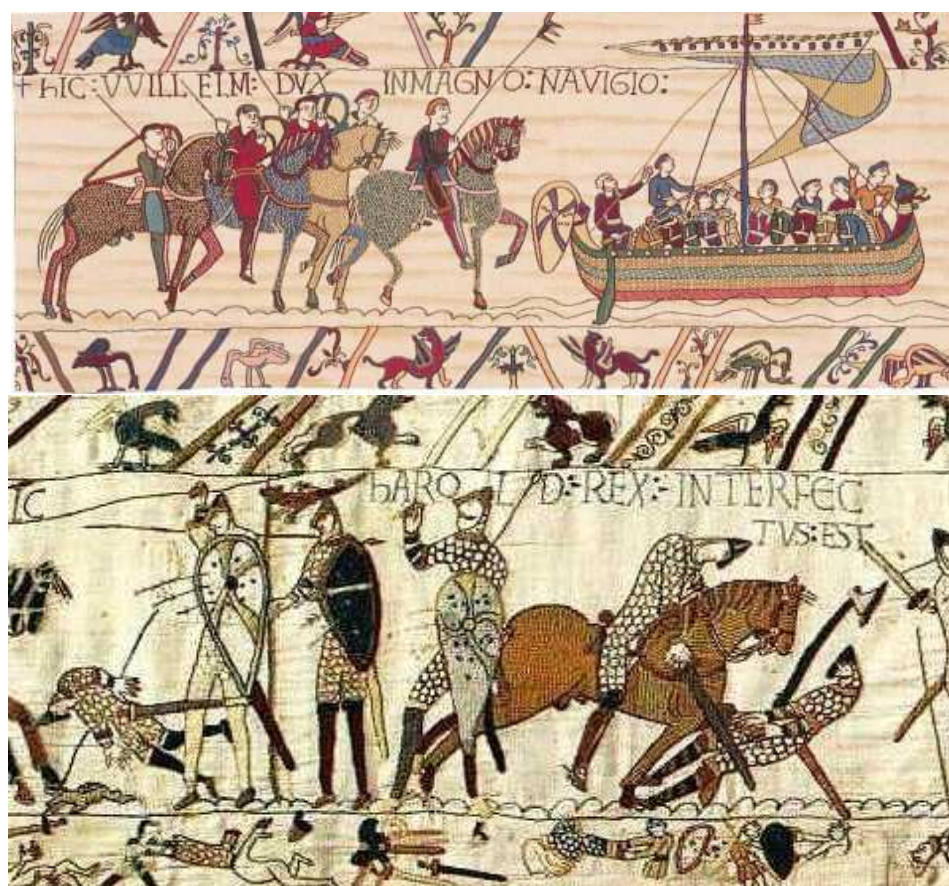


FIGURE 9.2 : Épisodes de la tapisserie de Bayeux montrant la Bataille d'Hasting en 1066. Cette tapisserie qui relate la victoire Normande et la conquête de l'Angleterre est un véritable exemple de propagande viking quant à leur férocité (<http://www.histoire-france.net>).

9.2.3 Témoignages écrits

Le code d'honneur verbal des Vikings a été très peu retranscrit en runes. Ainsi, ils ont laissé peu de témoignages écrits pouvant nous renseigner directement sur leur façon de vivre. Cependant, du fait qu'ils acceptèrent la christianisation autour du x^e siècle, avec une réticence variable selon les régions, leur société s'enrichit rapidement de lettrés. Ces personnes étaient principalement des ecclésiastiques et établirent divers registres, de mariages par exemple, nous permettant de dater certains événements historiques. De plus, autour du xii^e siècle, les épopées vikings furent retranscrites dans des manuscrits appelés sagas. Ces documents sont une source d'informations importantes car ils relatent bien plus que des histoires. En effet, ils contiennent de véritables biographies et généalogies qui peuvent dans une certaine mesure tenir lieu d'enregistrements historiques.

Une partie de notre argumentaire repose principalement sur les interprétations de quatre sagas :

Landnámabók saga écrite au XI^e siècle contenant la colonisation de l'Islande par les Scandinaves, l'auteur présumé se nomme Ari Þorgilsson.

Íslendingabók ou la saga des Islandais, qui concerne la colonisation et la généalogie des familles d'Islande, écrite au début du XII^e siècle.

Grœlendinga saga ou la Saga des Groenlandais. Cette saga relate l'histoire de la colonisation du Groenland et du Vinland (Canada), retranscrite au XIV^e siècle.

Eiríkr rauða saga ou la saga d'Erik le Rouge. Écrite probablement au XIII^e siècle et retranscrite au XV^e siècle, elle relate les colonisations du Groenland et de l'Amérique du Nord parfois dénommée saga de Thorfinn Karlsefni (*Þorfinnr Karlsefni*) véritable héros de cette saga [Babcock, 1917].

Écrits plusieurs siècles après les faits qu'ils relatent, ces manuscrits contiennent des informations qui se recroisent, attestant une certaine véracité. Par ailleurs, plusieurs personnages ayant réellement existé ont été identifiés par les historiens. Bien que l'authenticité de ces légendes, la généalogie ou les querelles qu'elles content puissent être remises en question, l'atmosphère, le mode de vie ainsi que le quotidien viking y sont décrits. Ces documents constituent en ce sens un outil précieux pour des recherches archéologiques. Par ailleurs, les sagas ont révélé aux historiens que les Groenlandais avaient accosté sur le nouveau continent près de cinq siècles avant Christophe Colomb. Ces documents sont à l'origine d'investigations archéologiques qui ont conduit, entre autres, à la découverte de l'Anse-aux-Meadows [Wallace, 2000]

9.2.4 Vestiges archéologiques

Au Groenland, les traces laissées dans le paysage par les Vikings témoignent de leur présence. Ils se sont établis sur des terres à l'époque inoccupées et ont construit des édifices en pierres caractéristiques qui n'ont pu être érigés par les populations voisines comme les Indiens d'Amérique du Nord ou les Inuits. En ce sens, tous les relevés archéologiques de bâtis antérieurs à 1721 (date d'arrivée du missionnaire luthérien, Hans Egede, au Groenland) ne peuvent être que vikings. Ceci confère une qualité remarquable à notre échantillon, car on ne peut absolument pas douter de l'attribution culturelle des occurrences utilisées pour la modélisation.

9.2.5 Le *Landnám* ou la colonisation des terres nordiques

La colonisation des régions arctiques et subarctiques par les Scandinaves s'effectue d'Est en Ouest sur plusieurs siècles et est connue sous le nom de *Landnám*. Partis des pays scandinaves, les Vikings ont découvert successivement les îles Féroé, l'Islande, le Groenland et le Canada. Le peuplement de ces régions est aujourd'hui un fait avéré et de nombreuses

preuves historiques attestent de son étendue. Nous pouvons grâce à des documents pour la plupart écrits par des religieux chargés de christianisation dater avec une précision relativement fine les événements de colonisation ainsi que les abandons de lieux. Il apparaît alors que cette conquête de l'Atlantique Nord fut soumise à deux épisodes climatiques, l'un favorable à la colonisation se déroulant pendant le OCM et l'autre correspondant au début de la détérioration climatique du PAG, climat moins approprié au mode de vie scandinave.

9.2.5.1 De la Terre de Glace à la Terre Verte

Avant l'arrivée des Norvégiens, l'Islande n'avait été peuplée que par quelques moines aux VII^e et VIII^e siècles, comme il est signalé dans le traité de Géographie *De mensura orbis terrae* de Dicuil [Tierney *et al.*, 1967 cités par Dugmore *et al.*, 2005]. C'est donc une terre inoccupée que découvrit Floke Vilgerdson vers 865 AD ; d'après la saga *Landnámabók*, ce Norvégien échoua dans un fjord rempli de glace et décida de nommer l'île *Ísland* qui signifie, en vieux norrois, terre de glace. Il faut attendre ensuite l'an 874 pour que débute le peuplement de l'île qui durera jusqu'en 930 [Dansgaard *et al.*, 1975]. L'île était très attractive et atteignit rapidement une population de grande taille estimée, selon les sources, de 40000 à 100000 habitants [Karlsson, 2000]. Une telle surpopulation à laquelle s'ajoute une série de catastrophes naturelles va entraîner famines et exodes.

Deux personnages Eiríkr Thorvaldson (*Eiríkr Þorvaldsson*) dit Erik le Rouge (*Eiríkr Rauði*) et son père Thorvald Ásvaldson (*Þorvaldr Ásvaldsson*), bannis d'Islande et de Norvège pour meurtre (*Landnámabók*) [Gad, 1970; Karlsson, 2000], profitent du contexte environnemental difficile pour enrôler des colons vers de nouvelles terres. Ils partirent vers l'Ouest car ils avaient connaissance d'îles récemment découvertes par Gunnbjörn Ulfsson (fin du IX^e siècle ou début du X^e siècle) et par Snaebjörn Galti (en 978 AD). Ainsi, Erik le Rouge et sa flotte naviguèrent vers l'ouest et découvrirent plusieurs nouvelles terres dont une que le chef viking nomma Groenland en 982 AD [Mills, 2003].

L'auteur de la *Saga des Groenlandais* émet quelques doutes quant au choix du nom qu'Erik le Rouge attribua à ce pays. Pour lui, ce n'était pas un choix anodin car Groenland, qui signifie Terre Verte, venait de cette façon en opposition à Terre de Glace. D'après l'auteur, Erik le Rouge louant les qualités de ce nouveau pays caressait l'espoir de le rendre attractif. D'après Dansgaard *et al.* [1975] les suspicions de l'écrivain seraient faussées par le climat qui régnait lors de l'écriture du manuscrit et montrent que les variations de températures indiquent bien une période chaude et donc une végétation potentiellement verdoyante lors de la découverte. La végétation du Groenland de l'époque n'était probablement pas beaucoup plus importante que celle que nous trouvons en ces lieux de nos jours (c'est-à-dire des pelouses rases et des arbrisseaux parsemés). Cependant, Massa [2010] souligne, citant Boyer [1992], que certains mois de l'année l'île peut apparaître verdoyante.

Qu'il fût vert ou non, le Groenland se trouva néanmoins être très attractif. Ce phénomène fut accentué d'une part par l'exode d'Irlande provoqué par les famines et d'autre part par une navigation facilitée. En effet, selon les chroniques de navigation des marchands islandais, la

présence amoindrie de glaces dérivantes durant l'OCM aurait facilité la circulation jusqu'à de très hautes latitudes [Koch, 1945 cité par Massa, 2010]. Erik le Rouge bénéficia de ce phénomène lorsque, accompagné de quelques familles scandinaves, il fonda vers 985 AD l'*Østerbygden* (ou colonie de l'Est, située à la pointe Sud du Groenland). Une deuxième colonie vit rapidement le jour, le *Vesterbygden* (ou colonie de l'Ouest, situé quelque 400 km au Nord de la première dans le fjord Godthaab, voir figure 9.3). Une troisième colonie est également mentionnée, *Mellemygden* (colonie du milieu, car entre les deux autres colonies) mais cette dernière peut être considérée comme une extension de l'*Østerbygden* [Massa, 2012]. Bien que le paysage fut principalement constitué d'un sol stérile car empierré et glacé, il existait des trouées de verdure dans les vallées des fjords où les agriculteurs nordiques purent s'installer. Ces trouées n'étaient pas homogènes et certaines, relativement plus luxuriantes que les autres, furent réclamées et accaparées par Erik le Rouge et les chefs pionniers du *Landnám* [McGovern *et al.*, 2006]. Cet auteur indique qu'en dépit de leur faible

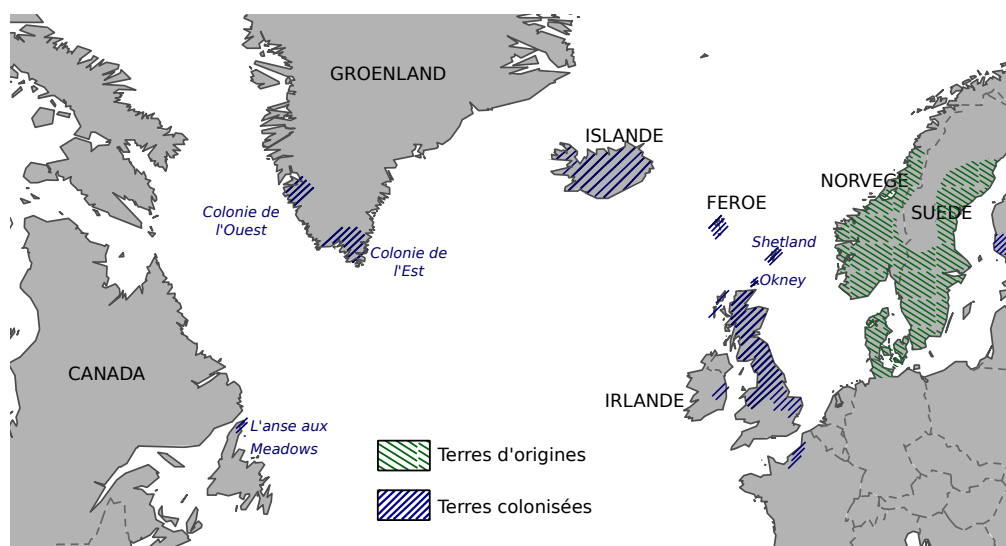


FIGURE 9.3 : Emprise des Scandinaves dans les régions arctiques, subarctiques, et dans le Nord de l'Europe, d'après Dugmore *et al.* [2005] (modifié).

potentiel agricole, on constate que la moindre parcelle de ces trouées a été occupée au moins par la première génération de colons. De cette manière la population du Groenland atteint 3000 à 6000 habitants (selon les sources [Lynnerup, 1998 cité par Dugmore *et al.*, 2005])

La soif de conquête des Groenlandais, ainsi que leur besoins en certaines ressources comme le bois (voir section 9.2.6) les poussèrent à embarquer pour de nouvelles terres plus à l'Ouest, jusqu'en Amérique.

9.2.5.2 Les Vikings en Amérique

Vers l'an mille, Leifr Ericson, fils d'Erik le Rouge lança une expédition vers l'Amérique. Les sagas la mentionnent et décrivent l'exploration de plusieurs territoires : le Markland (probablement Terre-Neuve) où se trouve le fameux site de l'Anse-aux-Meadows ; ainsi que d'autres régions comme l' Helluland (Île de Baffin) : et le Vinland dont la localisation fait débat, mais pourrait se situer aux alentours du Détroit de Cabot [Babcock, 1917]. La colonisation de l'Amérique a donc été initiée par les Vikings, mais n'a visiblement pas connu un essor comparable à celui des colonies du Groenland.

En Amérique comme au Groenland, les colons ont tenté de s'établir tout en conservant leur mode de vie, mais jusqu'à quel point cela a-t-il été possible ? Dans les prochains paragraphes nous tenterons de répondre à cette question en cherchant à définir tout d'abord sur quoi repose la subsistance de ces populations.

9.2.6 L'économie nordique

Des études et des simulations sur l'économie des fermes caractéristiques de l'époque viking d'Islande suggèrent que l'économie scandinave requérait une adaptation spécifique. On suppose qu'il en fût de même au Groenland [Adderley *et al.*, 2008; Keller, 1991] dans [Antunes *et al.*, 2014].

L'Islande et le Groenland se maintiennent dans un équilibre économiquement stable grâce à des échanges commerciaux. Au Groenland la production de produits arctiques comme l'ivoire de morse, les peaux d'ours polaires ou les faucons arctiques était échangée contre du fer, du sel, du bois, du miel ou des colorants. Le mode de subsistance des Scandinaves était principalement axé autour de l'élevage (bœufs, moutons, chèvres, cochons et chevaux) [Dugmore *et al.*, 2012]. Bien qu'ils aient cultivé pour leur propre consommation [Dugmore *et al.*, 2012; Fredskild et Humle, 1991], ils allouaient la majeure partie des surfaces arables à la production de fourrage pour alimenter le bétail. Ceci nécessitait des équipements spéciaux : champs clos, canaux d'irrigations et étables. Ces diverses constructions impliquaient une gestion du bois particulière et entraînèrent une déforestation [Adderley et Simpson, 2006; Adderley *et al.*, 2008; Arneborg, 2005; Buckland *et al.*, 2009; Dugmore *et al.*, 2012; Fredskild, 1992].

Nous avons vu précédemment que le bois était relativement rare dans le paysage groenlandais. Or, cette matière première était indispensable à leur subsistance. Leur seul recours était alors l'importation. La saga d'Erik le Rouge raconte que les Groenlandais allaient s'approvisionner en bois dans une terre dénommée *Markland*, signifiant « Terre des Forêts ».

McGovern *et al.* soutiennent que les pâturages limités à l'intérieur des fjords du Groenland n'ont probablement pas été suffisants pour reproduire le mode de subsistance islandais. Par conséquent, les Vikings durent se mettre à exploiter d'autres ressources comme les phoques ou les caribous [McGovern, 1985b]. Ces nouvelles pratiques les conduisirent à fréquenter une zone de chasse lointaine située à plusieurs centaines de kilomètres au nord

du *Vesterbygden* [McGovern, 1985a], aux conditions environnementales plus rigoureuses. Ces nouveaux comportements sont confirmés par les recherches archéozoologiques et isotopiques d'Arneborg *et al.* [1999] qui révèlent qu'au fil des siècles les Vikings du Groenland ont adopté, un tant soit peu, le mode de subsistance inuit en agrémentant des produits maritimes à leur alimentation [Arneborg *et al.*, 1999, 2012a,b]. De plus, l'ivoire ou les peaux des animaux chassés leur fournissait une valeur d'échange pour acquérir les matériaux dont ils manquaient [Frei *et al.*, 2015].

Le dernier enregistrement provenant du sud de l'*Østerbygden* (1407 AD), décrit deux activités : un mariage et la condamnation au bûcher d'une sorcière. En l'état actuel de nos connaissances, toute présence de Scandinaves au Groenland après cette date reste spéculative [Dugmore *et al.*, 2005]. Les fermiers scandinaves ont donc disparu du Groenland, aux alentours de cette date, la prochaine section évoque les différents facteurs ayant pu causer l'effondrement de cette société.

9.2.7 Déclin d'une civilisation

Les causes de la disparition des Vikings de ces régions est une question encore débattue de nos jours [Massa, 2012]. Des raisons d'ordre biologique, économique, politique ou écologique ont été évoquées (voir tableau 9.1 pour une liste non exhaustive). Les causes ayant provoquées la chute de cette société restent quelque peu incertaines. Cependant, les *scenarii* les plus cohérents incluent plusieurs de ces facteurs, souvent en connexion avec l'avènement du PAG. De fait, nous avons relevé dans la littérature une liste d'auteurs qui, sans pour autant exclure les raisons évoquées précédemment, soutiennent l'hypothèse que l'arrivée du PAG est une cause majeure de cette disparition [Antunes *et al.*, 2014; Barlow *et al.*, 1997; Dansgaard *et al.*, 1975; Diamond, 2005; D'Andrea *et al.*, 2011; Kuijpers *et al.*, 1999; Massa, 2012; McGovern, 1991; Patterson *et al.*, 2010; Stuiver *et al.*, 1995].

L'occupation du Groenland par les Scandinaves a duré au minimum 400 ans. Nous ne savons pas exactement ce qu'il a pu advenir de ces populations. Ont-elles périclité ou se sont-elles fondues dans les populations inuits ? Le XIV^e siècle, annonce la fin de la société groenlandaise. Le *Vesterbygden* disparaît peu avant 1350 AD, suivi une centaine d'années plus tard par de l'*Østerbygden* [Dansgaard *et al.*, 1975]. Cette disparition peut être le résultat d'une « intransigeance culturelle » c'est-à-dire l'obstination de la société à conserver son agriculture sans l'adapter aux nouvelles conditions climatiques [Barlow *et al.*, 1997; Diamond, 2005] ou d'une adaptation progressive les poussant à intégrer le mode de subsistance inuit mieux adapté au climat de l'époque en ces lieux [Arneborg *et al.*, 2012a; McGovern *et al.*, 2006].

Grâce aux reconstructions de niches écoculturelles, nous allons pouvoir mesurer l'influence du PAG sur l'effondrement de cette civilisation dans un premier temps, puis évaluer l'évolution de ces niches vikings dans le temps.

Tableau 9.1 : Liste non exhaustive des causes de la disparition des Vikings Groenland. Nombre de ces causes sont liées à la détérioration climatique du PAG.

Causes	Détails	Références
Inondations	La subsidence des terres sous le poids de la glace a pu causer une augmentation du niveau marin et inonder une partie des terres arables.	[Mikkelsen <i>et al.</i> , 2008]
Maladies	L'Europe et le Groenland connaissent de grandes épidémies.	[Barlow <i>et al.</i> , 1997]
Famines	La courbe des variations des températures suggère de nombreux épisodes de famine comme celui d'Islande en 1370.	[Dansgaard <i>et al.</i> , 1975]
Infertilité	Les populations isolées peuvent être victime d'infertilité congénitale.	[Barlow <i>et al.</i> , 1997]
Conflits	Les Scandinaves sont entré en contact avec les Amérindiens (ces derniers étaient dénommés <i>Skræling</i> dans la saga <i>Historia Norvegiae</i>).	[Barlow <i>et al.</i> , 1997]
Piraterie	Les contes des inuits blâment les pirates comme responsables de la disparition de cette société nordique affaiblie.	[Dansgaard <i>et al.</i> , 1975]
Isolement	La route maritime a été rompue à de multiples reprises à cause des intempéries et de l'apparition de glaces dérivantes, mais aussi pour des raisons politiques.	[Kuijpers <i>et al.</i> , 1999]
Sécheresse	Due aux vents, la sécheresse provoque des problèmes de gestion de l'eau douce pour l'irrigation.	[Lassen <i>et al.</i> , 2004]
Manque de fourrage	Une grande partie de l'économie de ces Vikings reposait sur l'élevage, la production de fourrage était donc capitale.	[Barlow <i>et al.</i> , 1997]
Manque de bois	Le bois était une ressource indispensable, utilisée pour la construction de granges (hivernage du bétail, stock du foin) ou de barrières (près clos).	[Barlow <i>et al.</i> , 1997]
Manque de Fer	La production d'outils en fer nécessitant beaucoup de bois, ces derniers devaient être importés.	[Barlow <i>et al.</i> , 1997]
Tempêtes	La période connut de forts changements de régimes des vents	[Kuijpers et Mikkelsen, 2009]
Anthropisation du milieu	Le défrichage et l'érosion du sol dus à l'introduction d'espèces domestiques non endémiques a pu provoquer la détérioration de l'environnement.	[Dugmore <i>et al.</i> , 2005; Massa <i>et al.</i> , 2012]

9.3 ECNM des Vikings des régions subarctiques

Dans ce travail les ECNM des Vikings des régions nordiques ont fait l'objet de deux études. La première a montré l'impact du PAG sur le peuplement du Groenland [Antunes *et al.*, 2014]. La seconde, utilisant des données environnementales plus précises, a confirmé les résultats de la première et a permis une interprétation plus avancée.

Comme nous l'avons vu, deux types de données sont nécessaires pour reconstruire les niches écoculturelles : des occurrences de présences archéologiques et des couches de données environnementales.

9.3.1 Données archéologiques

Nous avons utilisé pour les deux études les mêmes données archéologiques. Elles ont été extraites de cartes issues de la littérature [Adderley *et al.*, 2008; Krogh, 1982; McGovern

et al., 2007] selon la méthode décrite section 3.3 (figure 3.1) d'une part, de l'application Google Earth Visque du projet NABO (*North Atlantic Biocultural Organisation*) d'autre part.

Au Groenland les données archéologiques, essentiellement des localisations de fermes, sont d'une grande fiabilité car leur attribution culturelle ne peut pas être contestée. En effet, avant les Scandinaves, aucune population ne construisait dans ces régions de tels édifices. Après leur disparition, l'île ne fut peuplée que par des populations Inuits qui ne bâtaient rien en dur, du moins jusqu'au XVIII^e siècle (section 9.2.4). Les données étant d'une grande fiabilité, le biais d'échantillonnage est par conséquent réduit et nous n'aurons pas à appliquer de seuil d'omission (voir section 2.5.2). Les occurrences qui ont été collectées puis géoréférencées sont reportées sur la figure 9.4. On y distingue bien la séparation géographique des colonies *Østerbygden* et *Vesterbygden* présentées précédemment dans la section 9.2.5.1.

Nous avons choisi de séparer en deux lots ces occurrences conformément aux faits historiques afin de pouvoir apprécier s'il y a eu une évolution de niche écoculturelle. Si le peuplement du Groenland par les Scandinaves ne s'est fait qu'en une seule étape, son abandon quant à lui s'est produit en deux temps. En effet, Ivar Baardson, un émissaire épiscopal, signale l'abandon définitif du *Vesterbygden* dès l'an 1350 [Massa, 2012] alors qu'un acte de mariage célébré en 1408 a été retranscrit en 1409 en *Østerbygden* [Gad, 1970]. Ce document signale donc la présence de Vikings dans cette région au moins jusqu'à cette date.

Afin d'apprécier l'impact du changement climatique sur le maintien ou l'abandon des colonies vikings, nous avons séparé les sites en deux lots en fonction de ces événements. Le premier correspond aux XI^e - XIII^e siècles et le second aux XIV^e - XV^e siècles. Nous les appelons respectivement « 1^{er} peuplement » et « 2nd peuplement ». Sur la carte, figure 9.4, les sites qui ont disparu en premier, c'est-à-dire dès la fin de l'optimum climatique, sont représentés en orange et les sites qui ont résisté le plus longtemps à cette période froide sont représentés en bleu. Les sites du 1^{er} peuplement sont donc représentés par l'ensemble des points alors que les sites du 2nd peuplement sont figurés par les points bleus uniquement.

Nous nous sommes basés sur des datations C¹⁴ [McGovern *et al.*, 2007] pour l'affectation des sites islandais dans l'un ou l'autre des peuplements. Nous avons placé le site de l'Anse-aux-Meadows dans le 1^{er} peuplement car nous ne savons pas à quelle date il a été abandonné.

Dans l'étude de faisabilité, nous avons reconstitué les ECNM de chacun des peuplements.

9.3.2 Étude de faisabilité

Cette étude pilote a été publiée dans les actes du colloque *Debating Spatial Archaeology* [Antunes *et al.*, 2014].

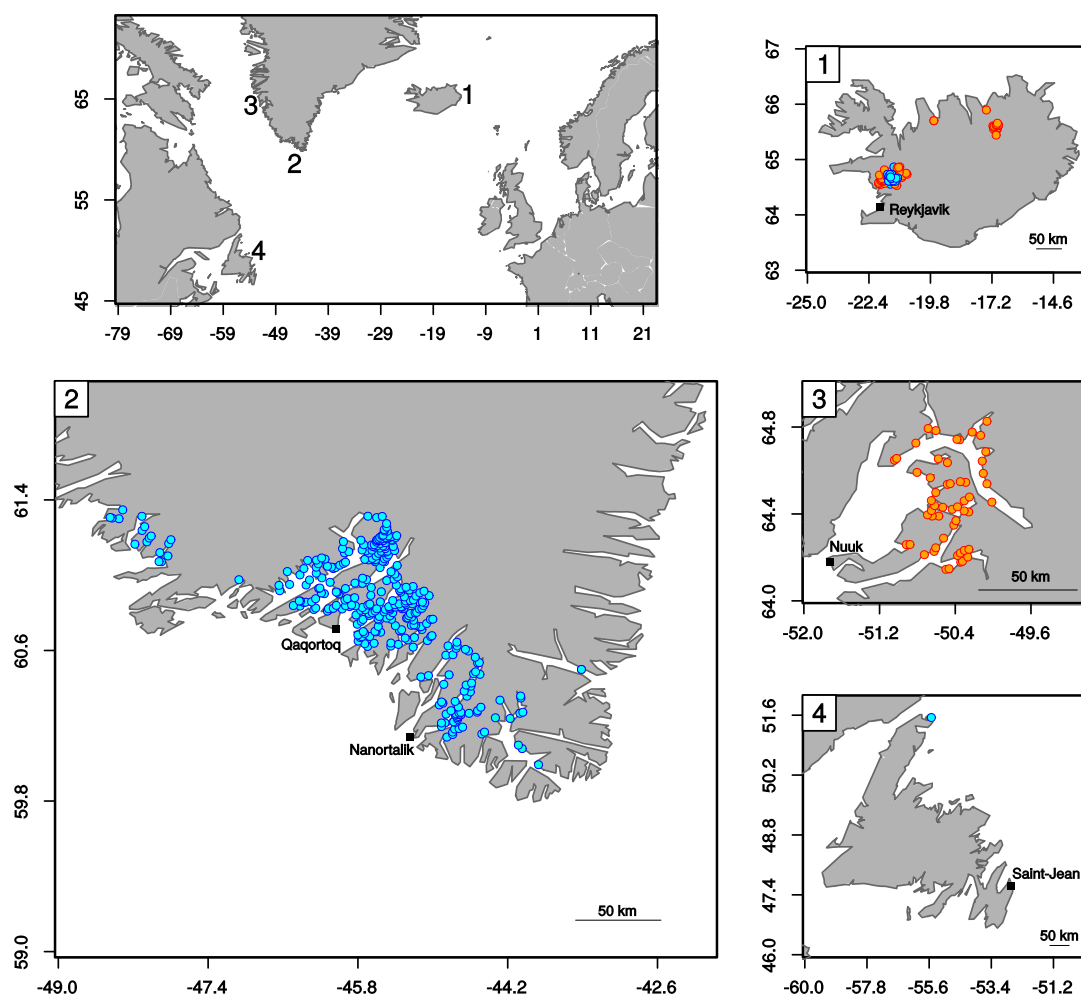


FIGURE 9.4 : Localisations des occurrences archéologiques vikings utilisées pour l'ECNM. Les points représentent les occurrences utilisées pour la modélisation de niches écoculturelles (fermes vikings) alors que les carrés indiquent la position de villes actuelles. Les occurrences archéologiques ont été recueillies pour quatre régions : l'Islande (1) ; l'*Østerbygden* et le *Mellebygden* (2) ; le *Vesterbygden* (3) ; et Terre-Neuve (4). Pendant le OCM Les Vikings ont occupé les fermes indiquées par les cercles oranges et bleus ($n = 559$). Les fermes occupées pendant le PAG sont représentées par les cercles bleus ($n = 450$).

9.3.2.1 Données environnementales

Comme cela peut souvent être le cas en ENM ou ECNM, nous sommes confrontés à un manque de données environnementales précises (voir section 2.3.3.1) pour la région et aux époques ciblées. Il n'existe en effet ni de zoom climatique ni de *downscaling* suffisamment précis pour l'étude que nous envisageons d'effectuer. Les simulations existantes possèdent une résolution très grossière et il est courant d'avoir un maillage pour ces latitudes de plus de 200 km. Une telle résolution ne permet pas une ECNM convenable d'autant plus que les distances maximales entre nos occurrences ne sont guère plus grandes (figure 9.4). Les données environnementales de cette étude proviennent de l'interpolation spatiale des variables radiations solaires de surface, température de l'air à 2 mètres, pression atmosphérique et précipitations de la simulation mensuelle de Jungclaus *et al.* [2010] qui couvre les deux derniers millénaires. L'interpolation de la simulation nous a permis le passage d'un maillage de 3.75° à 0.5° (environ 25 km à 65°N de latitude). L'interpolation consiste à découper chaque point de grille en plusieurs points de grille plus petits afin d'obtenir un maillage plus fin. Les valeurs des nouveaux pixels sont graduées et varient en fonction des valeurs de chaque pixel de la simulation basse définition. Cette interpolation a été réalisée grâce à la fonction *interpolate* de la bibliothèque *raster* [Hijmans et van Etten, 2012] de R.

Afin de refléter la saisonnalité et être en accord avec les pratiques agro-pastorales des Vikings, nous avons choisi d'inclure comme couches environnementales les variables pour les mois d'avril, de juillet et d'octobre, ainsi que les moyennes annuelles. Nous avons ajouté à ces 16 variables climatiques, l'altitude comme variable topographique (source : *Global Multi-resolution Terrain Elevation, Data 2010* - <http://eros.usgs.gov/>).

Une fois ces variables obtenues sur la période 800 - 2100 AD, nous avons procédé au découpage des périodes qui nous intéressaient et pour lesquelles nous avons calculé la moyenne de chacune des variables. Ainsi, nous obtenons des simulations climatiques à 0.5° pour les périodes 985-1250 AD et 1300-1450 AD que nous avons utilisées pour les ECNM des 1^{er} et 2nd peuplement vikings dans l'étude de faisabilité. Les niches écoculturelles ont été modélisées grâce au logiciel DesktopGARP (voir [Antunes *et al.*, 2014], pour les détails sur la méthode employée).

9.3.2.2 Principaux résultats de l'étude préliminaire

La projection géographique de la niche écoculturelle du 1^{er} peuplement durant l'optimum médiéval climatique (figure 9.5(a)) correspond à l'Islande dans sa quasi-totalité ainsi qu'au Sud et à l'Ouest du Groenland, soit dans les régions où s'étaient implantées les colonies de l'Est (*Østerbygden*) et de l'Ouest (*Vesterbygden*). On voit également une prédiction de présence faible, probabilité autour de 0.4, cette niche faiblement prédite s'étend dans la région du Labrador. Nous remarquons que ces résultats sont en accord avec les faits historiques. En effet, lorsqu'on confronte l'expression géographique des niches écoculturelles du 1^{er} peuplement et l'emprise maximale des Vikings (figure 9.3) dans ces régions

durant l'OCM, nous observons que ces derniers occupaient une grande partie de leur niche potentielle (section 1.4.2). Ceci signifie qu'ils s'étaient établis dans une grande partie des régions où l'environnement leur permettait de survivre avec leur mode de subsistance.

La reconstruction de la niche écologique des sites vikings après l'OCM (figure 9.5(b)) montre une réduction de surface par rapport à la prédiction précédente, ce qui se trouve être également en accord avec les faits historiques. On note particulièrement une forte baisse de la probabilité de présence dans la région du *Vesterbygden* ainsi qu'une diminution en intensité et en surface des zones prédites présentes dans la région de l'*Østerbygden*. La surface de la niche écoculturelle prédite du 1^{er} peuplement d'Islande est également plus large et plus forte que celle du 2nd peuplement. En ce qui concerne le Nouveau Monde, la prédiction change également. On note, pour le début du PAG, de fortes probabilités de présence à Terre-Neuve, au Nord de l'Anse-aux-Meadows et une disparition de la zone du Labrador où la présence potentielle avait été prédite comme faible pour la période précédente.

Pour conduire les modélisations des 1^{er} et 2nd peuplements, nous n'avons fait varier que les données climatiques et les occurrences. Les ECNM étant en accord avec les faits historiques, nous pouvons conclure que la répartition des Vikings dans ces régions était fortement liée au contexte climatique.

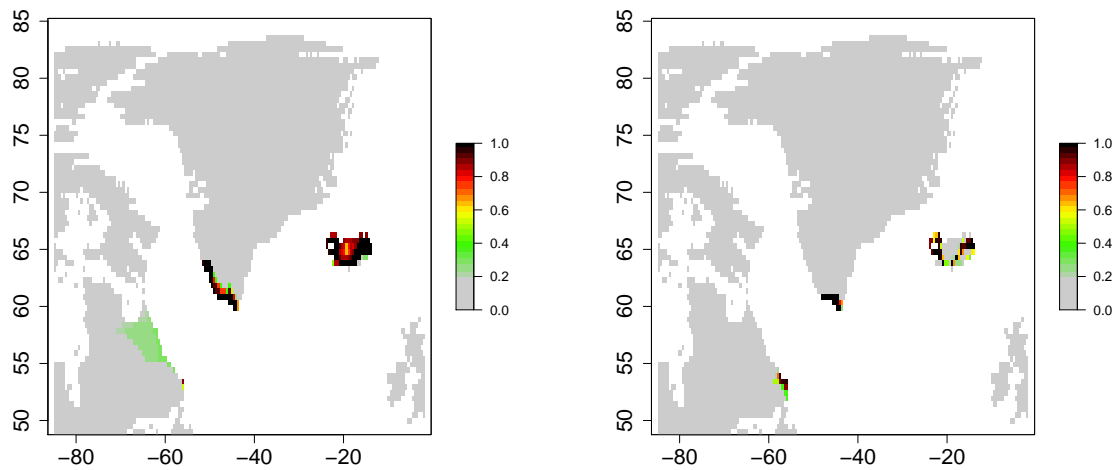
Les algorithmes prédictifs permettent de calculer une niche écoculturelle dans un environnement donné et de la projeter dans différents environnements. Dans la figure 9.5 sont représentés respectivement la projection de la niche écoculturelle du 1^{er} peuplement dans l'environnement du second (9.5(c)) et la projection de la niche écoculturelle du 2nd peuplement dans l'environnement du premier (9.5(d)). Dans chaque cas, on observe une réduction des aires de prédiction ainsi qu'une diminution d'intensité pour certaines régions prédites communes.

Le fait de ne pas observer une expansion lors de la projection à l'OCM du 2nd peuplement ni une prédiction très similaire à la niche prédite du 2nd peuplement lorsque l'on projette le 1^{er} peuplement dans le PAG signifie que les niches écoculturelles des deux peuplements sont différentes. Les projections nous permettent donc de constater que la niche écoculturelle des Vikings a été fortement modifiée suite au changement climatique.

Cette étude confirme bien l'impact important de la détérioration climatique sur le peuplement viking de ces régions arctiques. Toutefois la nature des variables, la résolution et la qualité de la simulation climatique sont ces biais importants qui ont pu nous conduire à de mauvaises conclusions. Afin de palier ce biais nous avons reconstruit de nouveau les niches écoculturelles des 1^{er} et 2nd peuplements en utilisant cette fois des données environnementales de meilleur qualité.

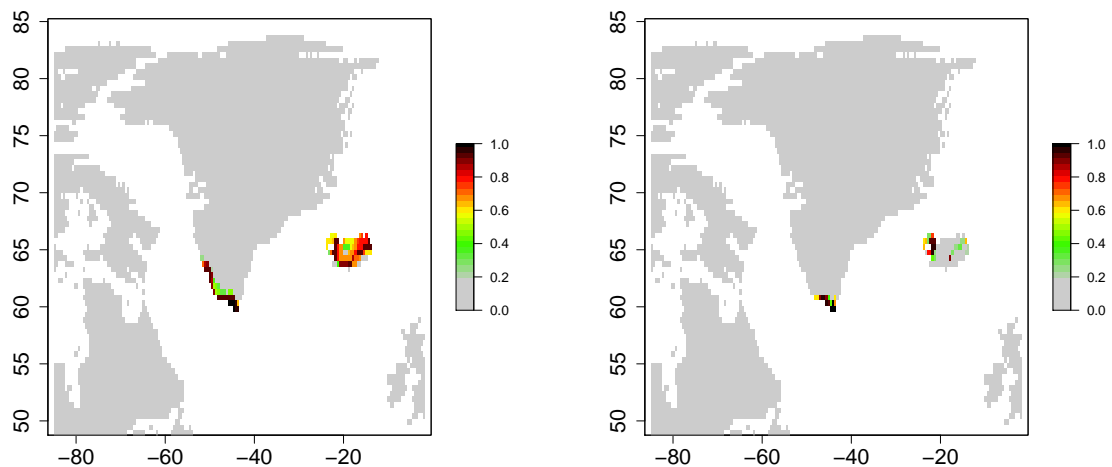
9.3.3 Une ECNM plus élaborée

L'utilisation de données environnementales de moindre qualité de l'étude de faisabilité font que la rigueur de la modélisation puisse être contestée. En effet l'interpolation spatiale



(a) Prédications de la niche viking du 1^{er} peuplement (période 985-1200 AD) avec l'algorithme DesktopGARP.

(b) Prédications de la niche viking du 2nd peuplement (période 1300-1450 AD) avec l'algorithme DesktopGARP.



(c) Projection de la niche viking modélisée avec DesktopGARP du 1^{er} peuplement dans l'environnement 1300-1450 AD.

(d) Projection de la niche viking modélisée avec DesktopGARP du 2nd peuplement dans l'environnement 985-1200 AD.

FIGURE 9.5 : Prédications des niches écoculturelles des Scandinaves obtenues en faisant la somme de 200 modèles réalisés avec DesktopGARP à partir d'une interpolation de la simulation climatique de [Jungclauss et al. \[2010\]](#).

d'une simulation climatique de basse résolution masque les variations climatiques qui se produisent à fine échelle, comme des variations dues à un ensemble de facteurs topographiques. Nous avons donc décidé de nous procurer des données à haute résolution pour les périodes concernées grâce à la méthode statistique de descente d'échelle présentée chapitre 4. Les données obtenues, de meilleure qualité, nous ont permis de réactualiser les ECNM des Vikings.

9.3.3.1 Optimisation des variables environnementales

La simulation de [Jungclauss *et al.*, 2010] ne possède pas les variables permettant d'obtenir des couches environnementales de type bioclimatique, soit les températures maximales, minimales et les précipitations (voir section 4.4.1.9). La première étape est donc de se procurer une autre source de données, c'est-à-dire une autre simulation climatique comprenant ces variables.

Il faut également veiller à ce que la simulation choisie soit en accord avec la réalité, ou ce que l'on connaît de la réalité, pour l'époque ciblée. En effet il existe des enregistrements de diverses natures prouvant les phénomènes climatiques de l'OCM et du PAG (section 9.1). Par conséquent, nous devons choisir une simulation climatique qui reflète la baisse de température après l'OCM, c'est-à-dire une simulation dont les températures varient conformément à la courbe d'anomalie établie par Mann (figure 9.1). La figure 9.6 montre les variations des températures moyennes annuelles de deux simulations, LMCMP5 [Swingedouw *et al.*, 2015] issue du modèle IPSL-CM5A-LR [Dufresne *et al.*, 2013] et SGI [Servonnat, 2010] issue du modèle IPSL-CM4_v2 [Marti *et al.*, 2005], au cours du dernier millénaire. La première simulation est réputée pour être généralement plus pertinente que la seconde. Cependant, au Groenland et pour la période sur laquelle nous focalisons notre attention (1000 - 1450 AD), elle reflète moins bien la chute des températures due à la sortie de l'OCM. Par conséquent, nous avons arrêté notre choix sur la simulation SGI pour construire les variables bioclimatiques à haute résolution nécessaires à la reconstruction des niches écoculturelles des Vikings.

Il s'agit d'une simulation mensuelle couvrant le dernier millénaire dans son intégralité. C'est une simulation globale avec une grille de 96×72 , soit une résolution de $3.75 \times 2.5^\circ$, ce qui équivaut à une grille d'environ 180×280 km à la latitude 65°N . Elle a été effectuée en considérant l'évolution des paramètres orbitaux, avec un forçage des gaz à effet de serre ainsi qu'un forçage solaire important. Ce dernier point vient appuyer notre choix car, selon [Mann *et al.*, 2009; Shindell *et al.*, 2001 cités par Servonnat, 2010], l'activité solaire serait à l'origine des différences d'indices de l'oscillation nord atlantique entre l'OCM et le PAG.

L'application de la méthode de *downscaling* présentée dans le chapitre 4 a permis l'obtention à partir de la simulation SGI de couches environnementales bioclimatiques de résolution 10' (environ 8 km) pour les périodes 1000-1300 AD et 1301-1450 AD ainsi que pour chaque siècle du x^e au xv^e siècle. Nous avons utilisé ces couches environnementales pour la nouvelle ECNM.

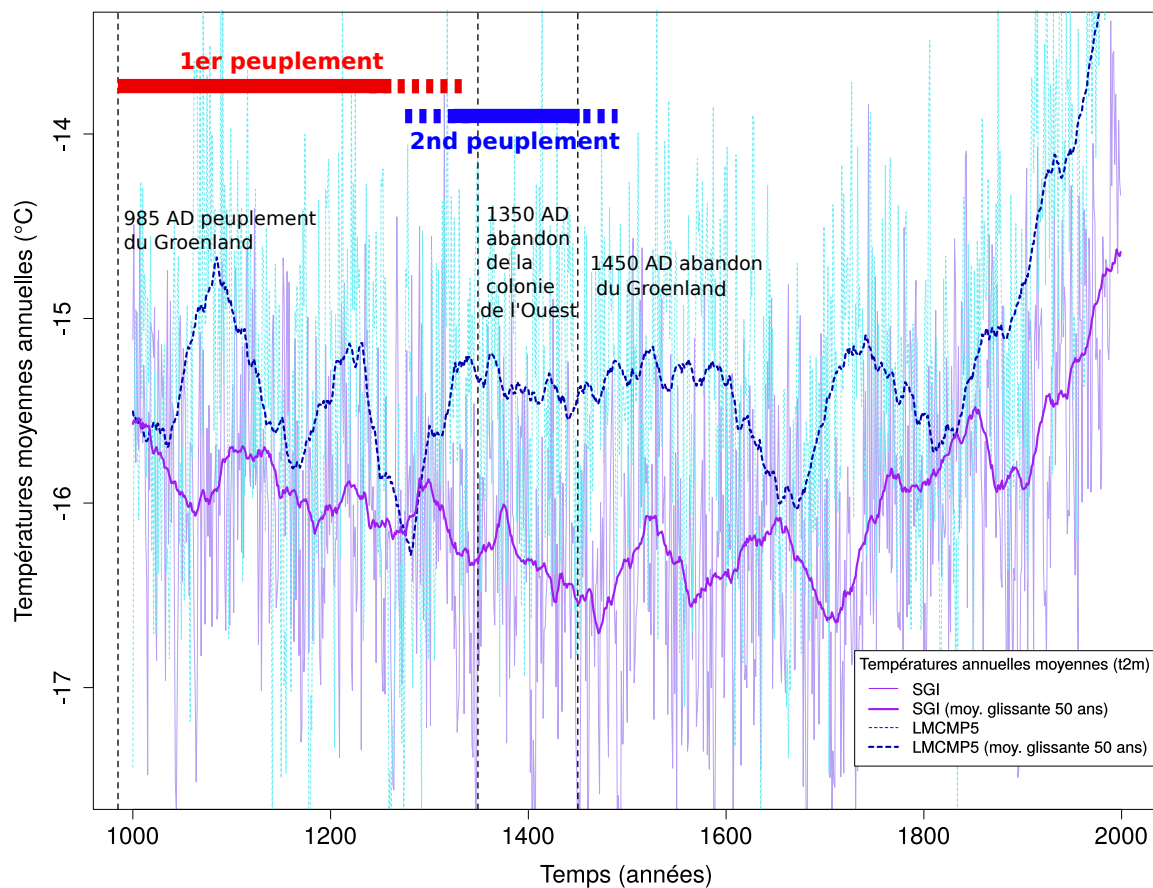


FIGURE 9.6 : Courbes de variations des températures moyennes annuelles au Groenland selon la simulation LMCMP5 [Swingedouw *et al.*, 2015] et la simulation SGI [Servonnat, 2010] au cours du dernier millénaire. Sont représentées les moyennes annuelles brutes (traits fins) ainsi que les moyennes glissantes selon une fenêtre de 50 ans (traits épais). Trois événements majeurs de l'histoire des colonies scandinaves du Groenland sont indiqués : la fondation de l'*Østerbygden* par Erik le Rouge (985 AD), le début du dépeuplement du *Vesterbygden* (vers 1300 AD) et l'abandon définitif du Groenland (après 1450 AD).

Tableau 9.2 : Liste des algorithmes utilisés pour la reconstruction des niches écoculturelles des Vikings. Pour la réalisation des consensus élitistes pondérés, les algorithmes ont été sélectionnés en fonction de l'AUC (> 0.8) puis pondérés en fonction du *partial ROC ratio* (proc). Douze algorithmes ont été utilisés avec leurs réglages par défaut plus les algorithmes DesktopGARP et MaxEnt avec des réglages leur donnant moins de restrictions dans leurs prédictions (optimistes).

	1 ^{er} peuplement				2 nd peuplement			
	Préd. 1000-1300		Proj. 1301-1450		Préd. 1300-1301		Proj. 1301-1450	
	auc	proc	auc	proc	auc	proc	auc	proc
Bioclim	0.996	1.000	0.996	1.000	0.500	1.000	0.500	1.000
Chebyshev Env. dist.	0.999	1.997	0.999	1.998	0.999	1.999	0.999	1.998
DesktopGARP	0.500	1.000	0.500	1.000	0.500	1.000	0.500	1.000
DesktopGARP BS	0.500	1.000	0.500	1.000	0.500	1.000	0.460	1.000
DesktopGARP BS (optimiste)	0.269	1.000	0.278	1.000	0.132	1.000	0.437	1.000
DesktopGARP	0.191	1.000	0.500	1.000	0.441	1.000	0.500	1.000
Envelope Score	0.996	1.992	0.996	1.993	0.989	1.978	0.990	1.979
Environmental distance	0.500	1.000	0.500	1.000	0.500	1.000	0.500	1.000
Euclidean Env. dist.	0.500	1.000	0.500	1.000	0.500	1.000	0.500	1.000
MaxEnt	0.998	1.996	0.998	1.997	0.998	1.996	0.997	1.994
MaxEnt (optimiste)	0.500	1.000	0.500	1.000	0.500	1.000	0.500	1.000
Niche Mosaic	0.993	1.000	0.994	1.000	0.500	1.000	0.500	1.000
Random Forests	0.979	1.958	0.969	1.939	0.564	1.128	0.569	1.139
SVM	0.982	1.000	0.987	1.000	0.988	1.000	0.989	1.000

9.3.3.2 Optimisation des occurrences

Afin de modéliser les niches des 1^{er} et 2nd peuplements nous conservons les données d'occurrences archéologiques présentées section 9.3.1 et figure 9.4 Bien que nous soyons certains des attributions culturelles et chronologiques de ces occurrences, elles sont suffisamment nombreuses pour attribuer un taux d'omission de 5% qui permet d'optimiser la modélisation par l'application d'un seuil selon la méthode *Fixed sensitivity* décrite tableau 2.5.3.

9.3.3.3 Optimisation de la modélisation

Contrairement à l'étude de faisabilité pour laquelle nous avons réalisé les modèles à l'aide d'un seul algorithme prédictif (DesktopGARP), nous nous servons d'un panel d'algorithmes prédictifs afin de créer des consensus élitistes pondérés qui permettent l'obtention de résultats plus robustes suivant la méthode décrite chapitre 6.

Le tableau 9.2 dresse la liste des algorithmes utilisés pour cette étude ainsi que la pondération des prédictions qui ont servi à la réalisation du consensus. N'ont été retenus que les prédictions ayant des valeurs d'AUC supérieure à 0.8. Ensuite, ces dernières ont été pondérées par les valeurs correspondantes à leur *partial ROC ratio*.

Dans un premier temps, la modélisation des niches sur les périodes 1000-1300 AD et 1301-1450 AD nous permet de dresser leurs caractéristiques (étendue géographique, *breadth*, liste des principaux facteurs environnementaux) et d'apprécier leurs dynamiques

évolutives. Dans un second temps, nous effectuons des projections de la niche reconstruite pour la première période dans la seconde, puis inversement. Enfin, dans un troisième temps, nous menons une analyse avec un découpage temporel plus fin en prédisant les niches des XI^e et XIV^e siècles puis en les projetant séculairement sur toute la période étudiée afin d'évaluer leurs évolutions.

9.3.3.4 Analyse des niches écoculturelles vikings

La projection géographique de la prédiction de niche du 1^{er} peuplement (figure 9.7) est comparable aux résultats obtenus lors de l'étude de faisabilité avec cependant de forts contrastes concernant l'Islande et le Sud du Groenland. En effet l'affinement de l'échelle de la nouvelle étude permet d'observer des nuances de probabilités de présence non perçues par l'étude préliminaire. On constate que les prédictions de présence ont une probabilité de valeur moyenne pour la majeure partie de l'Islande (excepté des zones correspondant à la présence de glaciers où la probabilité de présence est nulle). La niche est prédite avec une forte probabilité dans deux zones. Une située au Nord de Reykjavik et l'autre aux alentours de Húsavík (Nord-Ouest). Ces « îlots » de prédictions de présence coïncident avec les occurrences vikings utilisées pour la modélisation de niches. Une variation entre les deux études est également constatée pour la région comprise entre les colonies du Groenland. Cette dernière montre des probabilités de présence fortes continues dans la première étude alors que la seconde étude y indique une prédiction plus faible. La meilleure résolution de la nouvelle modélisation permet également de distinguer une mosaïque de prédictions de moyennes probabilités de présence en Terre-Neuve où l'autre étude prédisait la zone entière.

La prédiction de la niche du 2nd peuplement (figure 9.9) signale une quasi-absence de présence en Islande alors que l'étude menée avec la simulation climatique interpolée y indique, pour la même période, des zones à fortes probabilités de présence. Il en est de même pour le Labrador où la nouvelle étude prédit une présence moyenne dans toute la région, à l'exception de l'Anse-aux-Meadows qui est prédite avec une forte probabilité. Les deux études sont comparables pour le Sud du Groenland, mais la fine résolution de la seconde modélisation permet de nouveau de distinguer d'une zone à probabilité de présence moins élevée entre les colonies séparant l'*Østerbygden* du *Mellemygden*. On note une probabilité de présence très faible aux alentours du *Vesterbygden* alors que cette colonie ne faisait pas partie de la niche prédite pour le 2nd peuplement dans l'étude pilote.

À l'instar de l'étude de faisabilité, nous avons projeté les niches prédites pour chacun des peuplements dans la période qui leur est postérieure ou antérieure. Ainsi, la niche prédite pour le 1^{er} peuplement est projetée dans l'environnement de la période 1301-1450 AD (figure 9.9) et la niche prédite pour le 2nd peuplement dans l'environnement de la période 1301-1450 AD (figure 9.10).

Dans les deux cas, nous observons une atténuation des probabilités de présence par rapport aux prédictions. Les conditions environnementales n'étant pas les mêmes, ce résultat

suggère une modification de niche. S'il y avait eu conservation de niche, la projection de la niche du 1^{er} peuplement dans la deuxième période aurait une expression géographique de surface différente, plus réduite. La faible variation entre niche prédite et niche projetée du 2nd peuplement amène à la même conclusion. En effet, la représentation géographique des niches indique également que le 2nd peuplement n'aurait pas occupé les mêmes localités que le 1^{er} peuplement s'il avait bénéficié du même environnement. Ces résultats mettent donc en évidence une réduction de niche. Afin de vérifier cette observation, nous allons procéder à l'analyse en composante principale des niches modélisées.

9.3.3.5 Évolution des niches / impact du climat

Comme nous l'avons vu section 7.4, les ACP permettent d'apprécier la répartition des lieux convenables aux niches prédites dans l'espace écologique. Nous avons réalisé quatre ACP individuelles correspondant aux prédictions et projections de chacun des peuplements ainsi qu'une ACP commune (section 7.4.1) qui regroupe les quatre niches modélisées. Cette dernière permet d'observer la position des niches entre elles dans l'espace écologique. Ces ACP sont élaborées à partir des relevés des valeurs de chaque variable environnementale pour chaque point des prédictions ou projections pour lesquels la modélisation a prédit une présence potentielle avec une probabilité supérieure à 0.8. Le plan de l'ACP représente alors l'espace écologique ou environnemental. Dans l'ACP commune, ce plan équivaut donc à l'union des espaces environnementaux des périodes étudiées (Env. 1000-1300 AD \cup Env. 1301-1450 AD). Les ellipses de dispersion de l'ACP représentent la projection moyenne de l'inertie des individus de l'ACP. Les axes des ellipses sont construits selon un facteur k égal à 1.5 fois l'écart-type des coordonnées de chaque point. Les ellipses synthétisent ainsi environ 67% de l'information apportée par les individus de l'ACP [Dray *et al.*, 2007].

Les analyses révèlent (figures 9.11 et 9.12) que les niches prédites pour chacun des deux peuplements ne sont pas similaires. La niche du premier est très large relativement à celle du second. De plus, la niche correspondant à la population de la période chaude englobe totalement la niche reconstruite pour l'autre population. Ceci suggère une restriction de niche. Lorsque nous observons la répartition des individus de l'ACP commune (figure 9.11), nous constatons que le 1^{er} peuplement est constitué de trois groupes qui se distinguent selon les conditions environnementales et qui correspondent à leurs localisations géographiques. On observe ainsi des amas de points pour l'Islande ainsi que pour chaque colonie du Groenland. Nous notons la position quasi centrale de l'Anse-aux-Meadows (site viking du Labrador) et remarquons également un décalage des centres des ellipses (reflétant l'optimum théorique des niches) qui suggère, en plus de la restriction, un décalage des niches suivant les deux axes et plus particulièrement sous l'influence des variables bio12 (précipitations annuelles), bio13 (précipitations du mois le plus humide) et bio16 (précipitations du trimestre le plus humide).

Les ACP individuelles nous permettent de définir les principaux facteurs environnementaux qui déterminent chaque niche (prédictions et projections). Peu de variations entre les

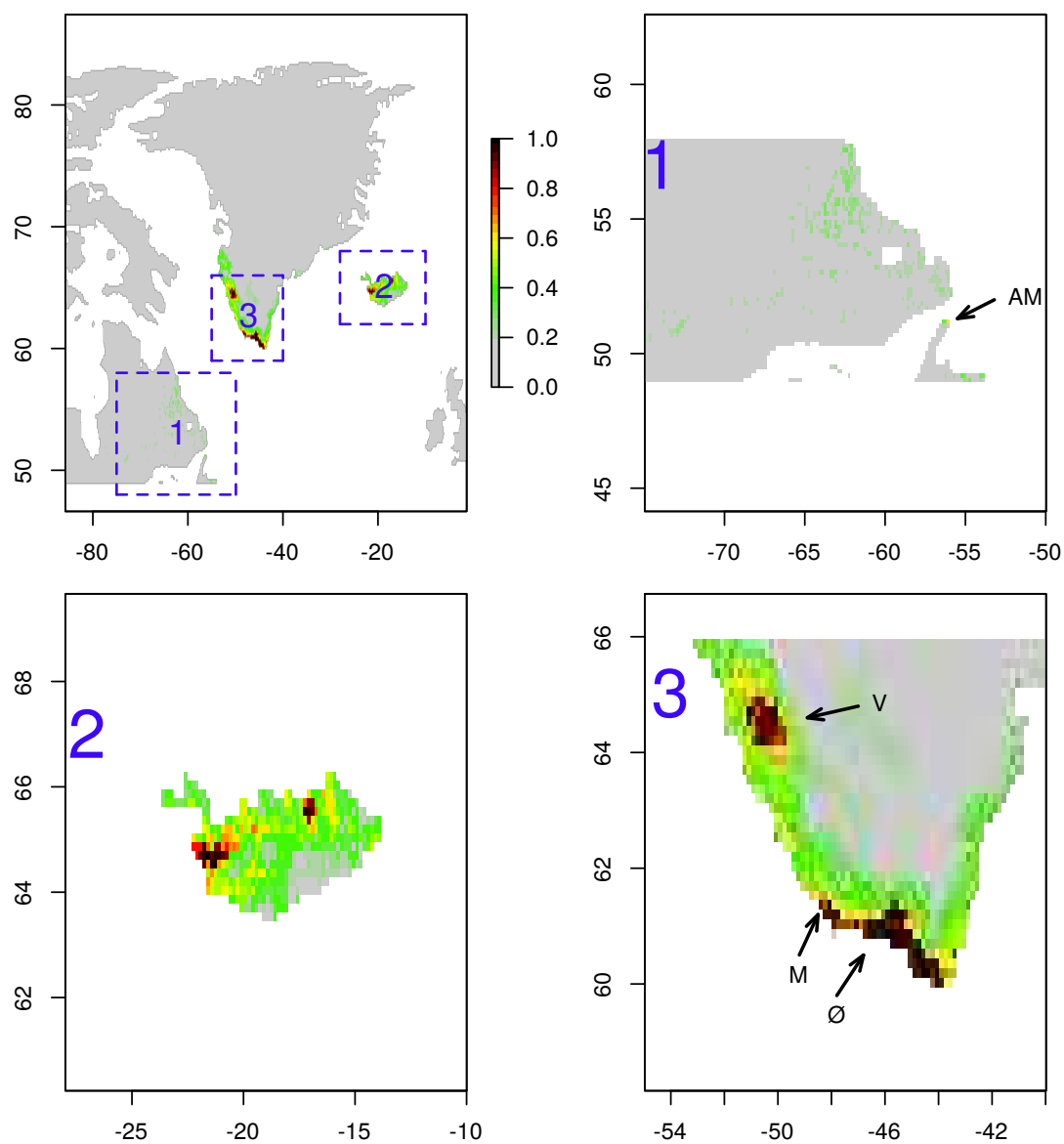


FIGURE 9.7 : Prédiction consensuelle (selon 14 algorithmes prédictifs) de la niche écoculturelle du 1^{er} peuplement viking durant l’OCM (environement 1000-1300) – (1) Labrador, (2) Islande, (3) Groenland, (AM) Anse-aux-Meadows, (Ø) Østerbygden, (M) Mellebygden, (V) Vesterbygden.

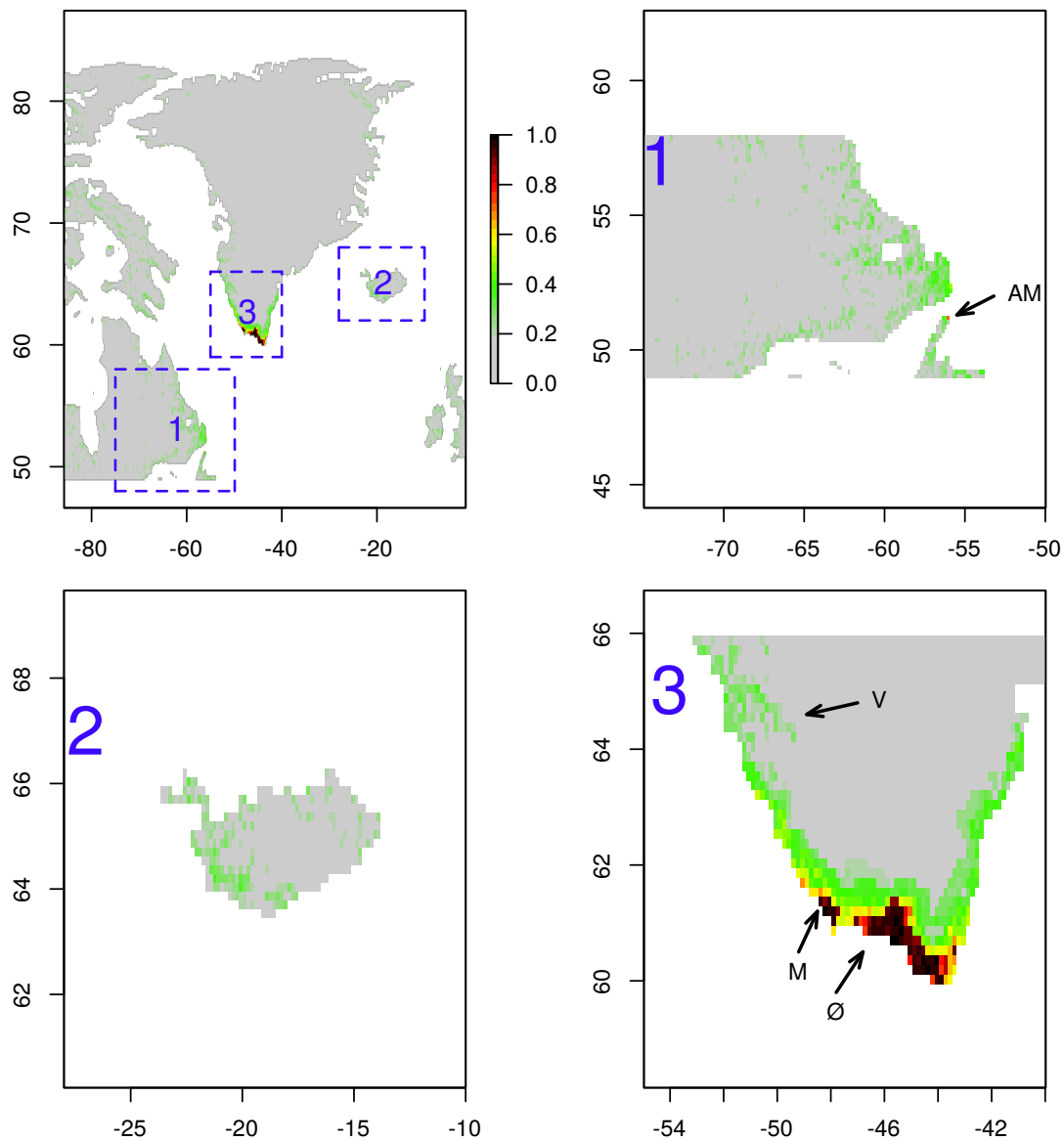


FIGURE 9.8 : Prédiction consensuelle (selon 14 algorithmes prédictifs) de la niche écoculturelle du 2nd peuplement viking durant la période 1301-1450 AD – (1) Labrador, (2) Islande, (3) Groenland, (AM) Anse-aux-Meadows, (Ø) Østerbygden, (M) Mellebygden, (V) Vesterbygden.

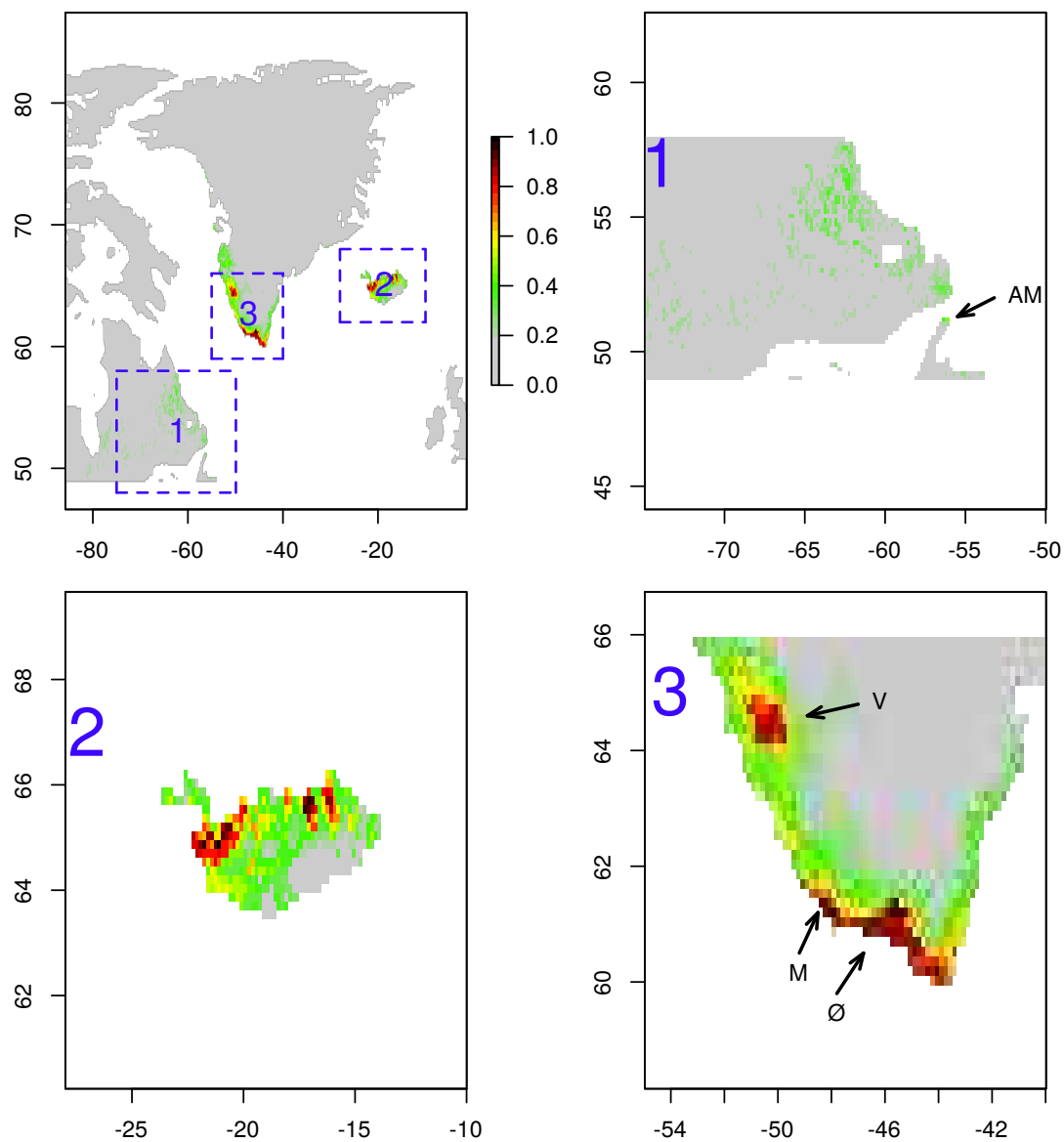


FIGURE 9.9 : Projection consensuelle (selon 14 algorithmes prédictifs) de la niche écoculturelle du 1^{er} peuplement viking dans l'environnement de la période 1301-1450 AD– (1) Labrador, (2) Islande, (3) Groenland, (AM) Anse-aux-Meadows, (Ø) Østerbygden, (M) Mellemygden, (V) Vesterbygden.

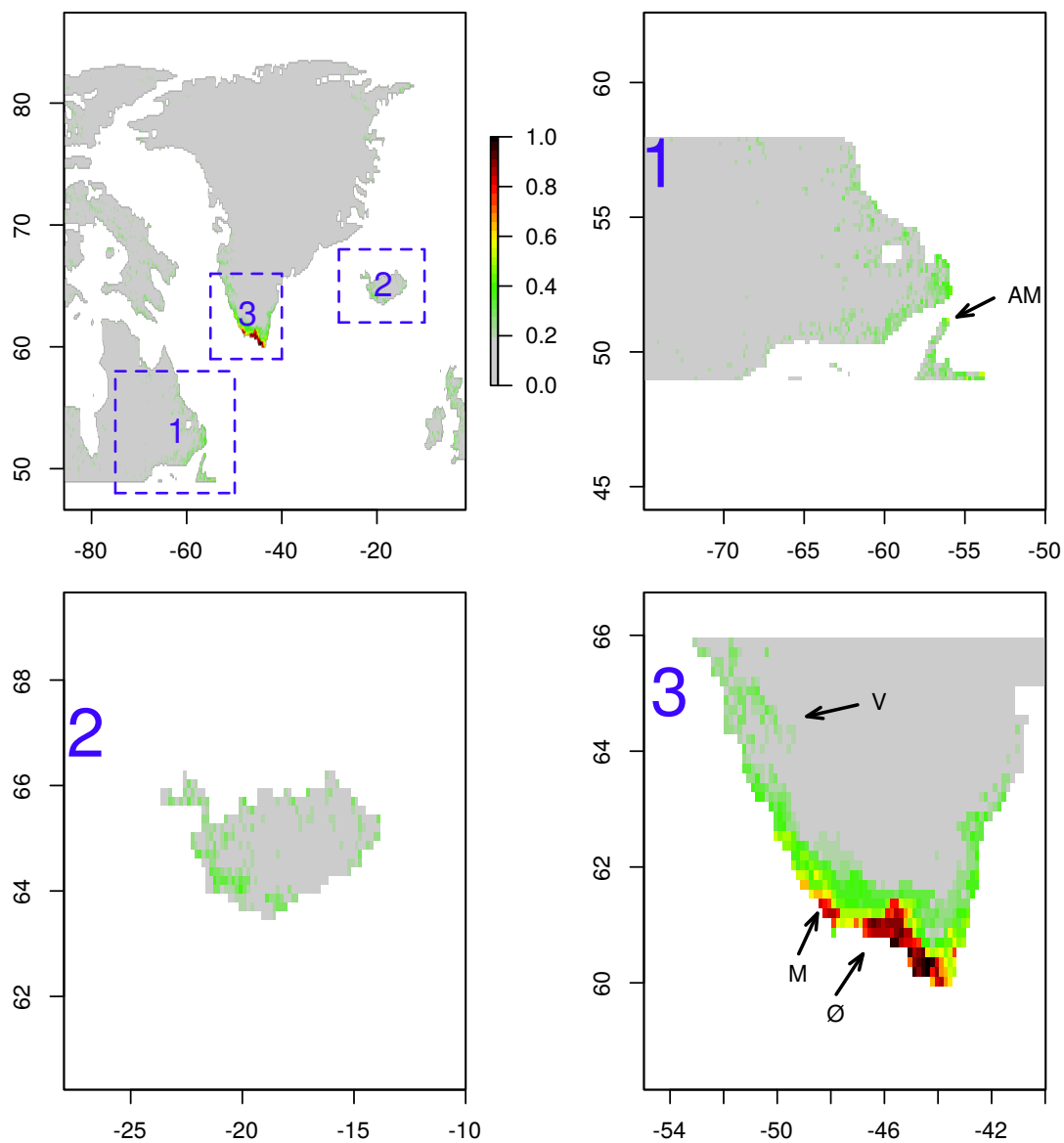


FIGURE 9.10 : Projection consensuelle (selon 14 algorithmes prédictifs) de la niche écoculturelle du 2nd peuplement viking dans l'environnement de l'OCM (1000-1300 AD) – (1) Labrador, (2) Islande, (3) Groenland, (AM) Anse-aux-Meadows, (Ø) Østerbygden, (M) Mellemygden, (V) Vesterbygden.

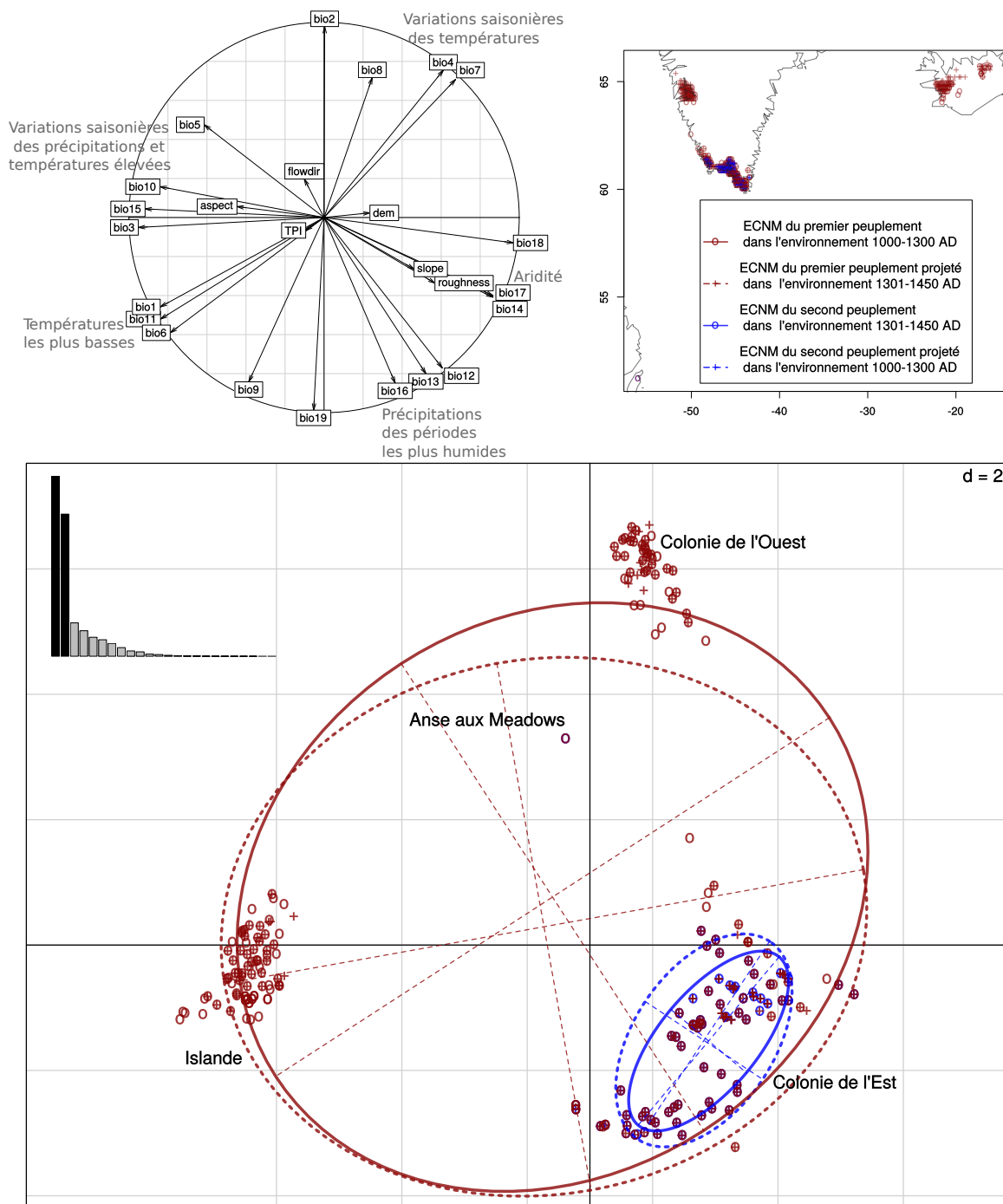


FIGURE 9.11 : Analyse en composantes principales commune des ECNM du Groenland. En haut à gauche est représenté le cercle de corrélation des variables ; en haut à droite les localisations géographiques des prédictions de présence utilisées pour l'ACP (probabilité de présence > 0.8). La partie inférieure de la figure représente les positions des prédictions de présence sélectionnées dans l'espace environnemental représenté par le plan de l'ACP dont le premier axe a une inertie de 39.69% et le second de 31.37%. Les ellipses synthétisent les barycentres des niches.

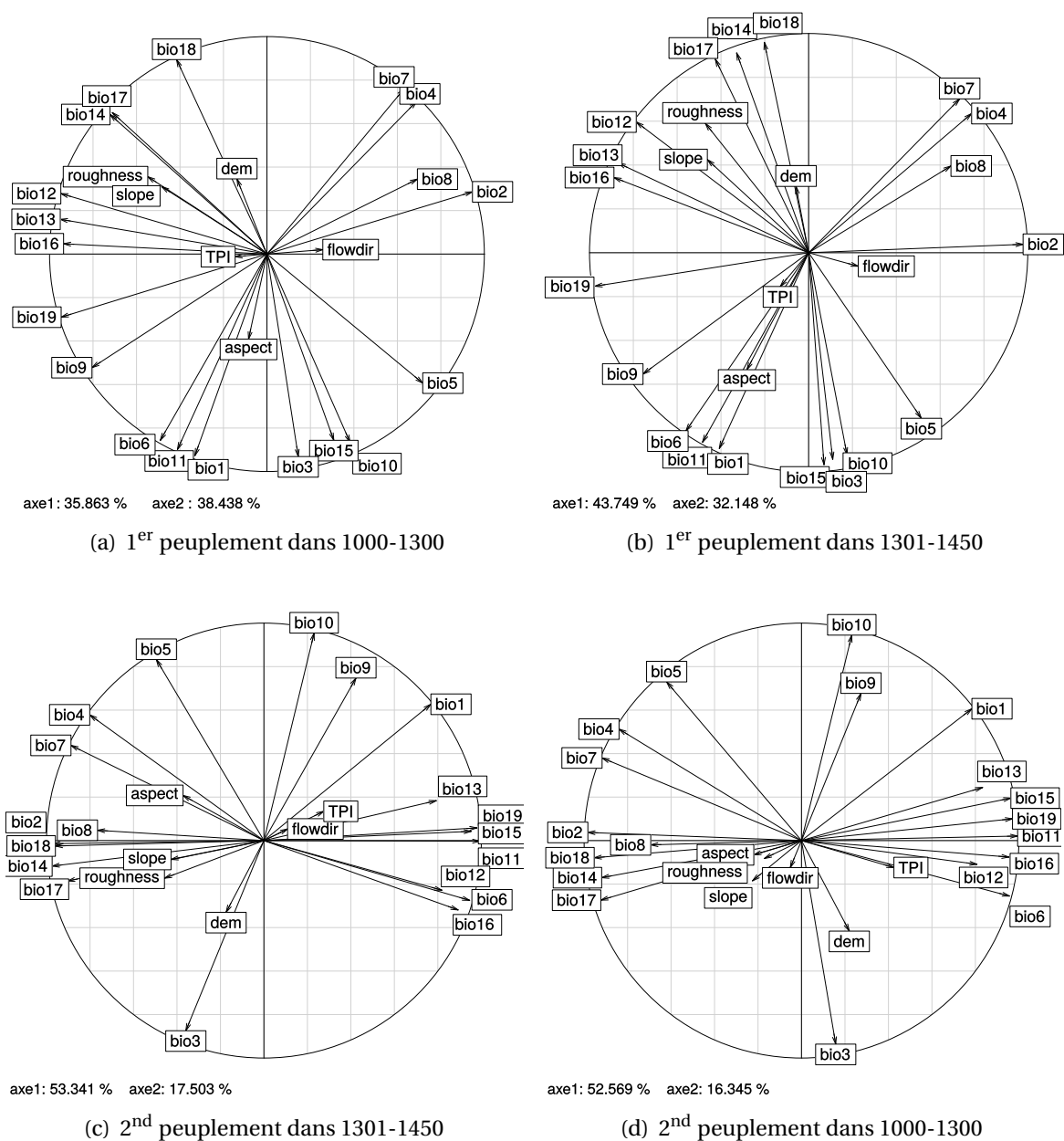


FIGURE 9.12 : Cercles de corrélation des variables environnementales (bioclimatiques et topographiques) des analyses en composantes principales des projections et prédictions des niches écoculturelles des peuplements vikings. L'axe 1 est l'axe des abscisses et l'axe 2 celui des ordonnées.

niches prédites et leur projections sont observables (dans la figure 9.12). Cela signifie que les expressions géographiques des niches se caractérisent par des combinaisons de conditions environnementales proches. On note toutefois des différences dans les contributions des variables topographiques des prédictions et des projections (figure 9.13) mais ces dernières ne contribuent que secondairement à l'élaboration de la niche. Les variables bioclimatiques

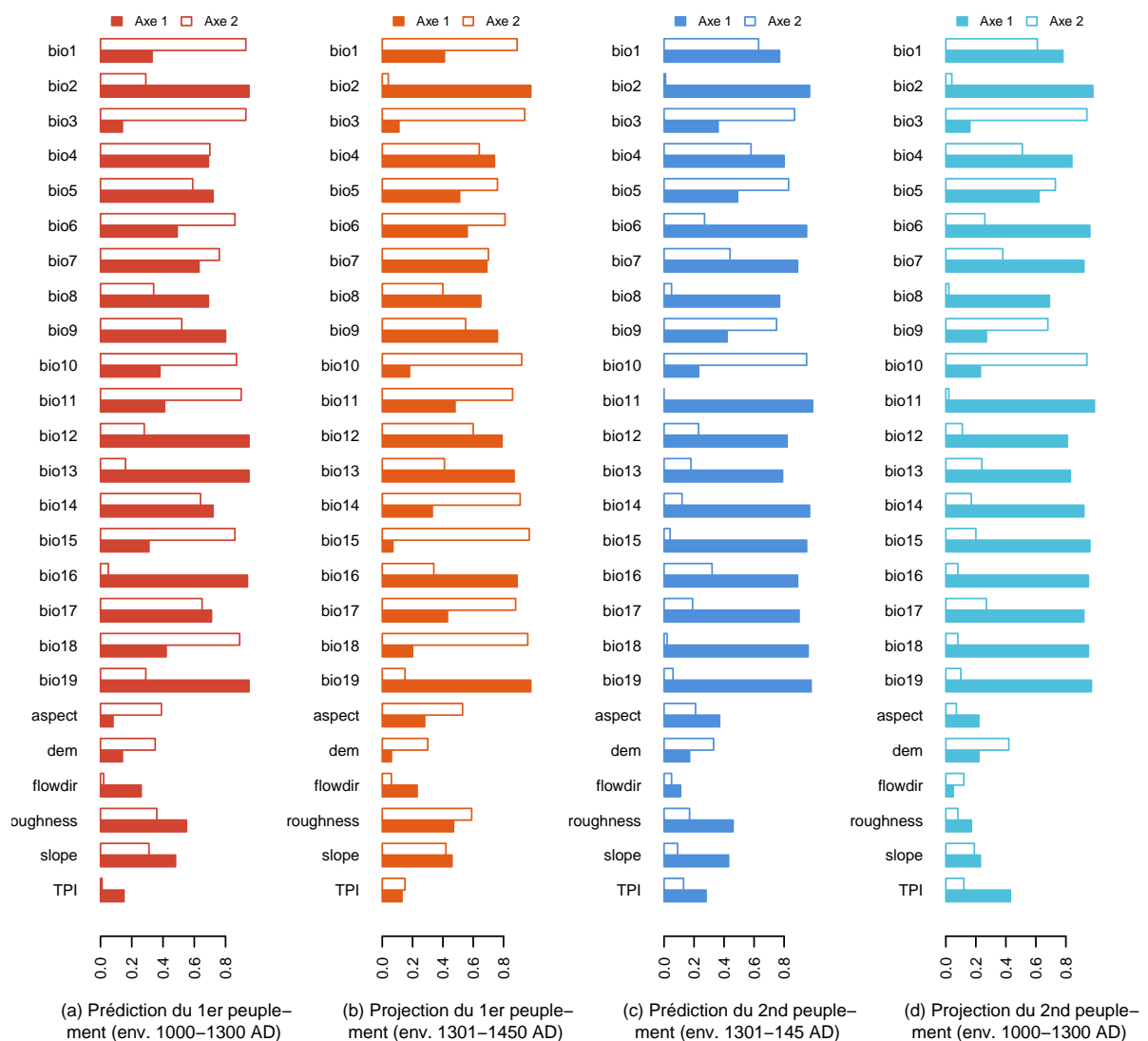


FIGURE 9.13 : Importance des variables environnementales

conservent, quant à elles, leurs relations entre les prédictions et projections. De plus ces dernières ont, selon leurs valeurs de première composante, les rôles les plus importants dans la structure des niches. Ce phénomène est également observable dans les figures 9.14 et 9.15 qui représentent les distributions des valeurs de chacune des variables dans les niches modélisées.

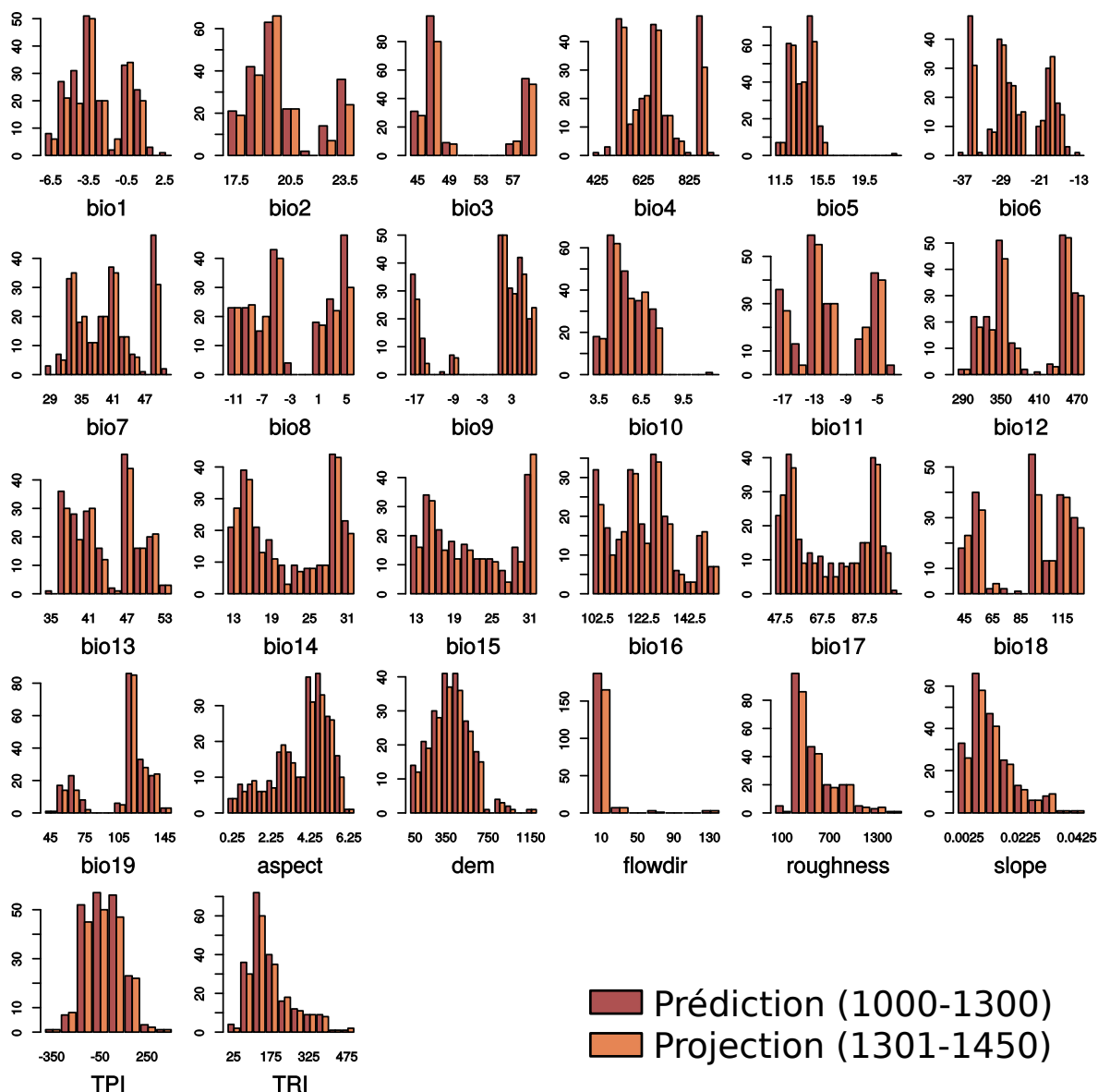
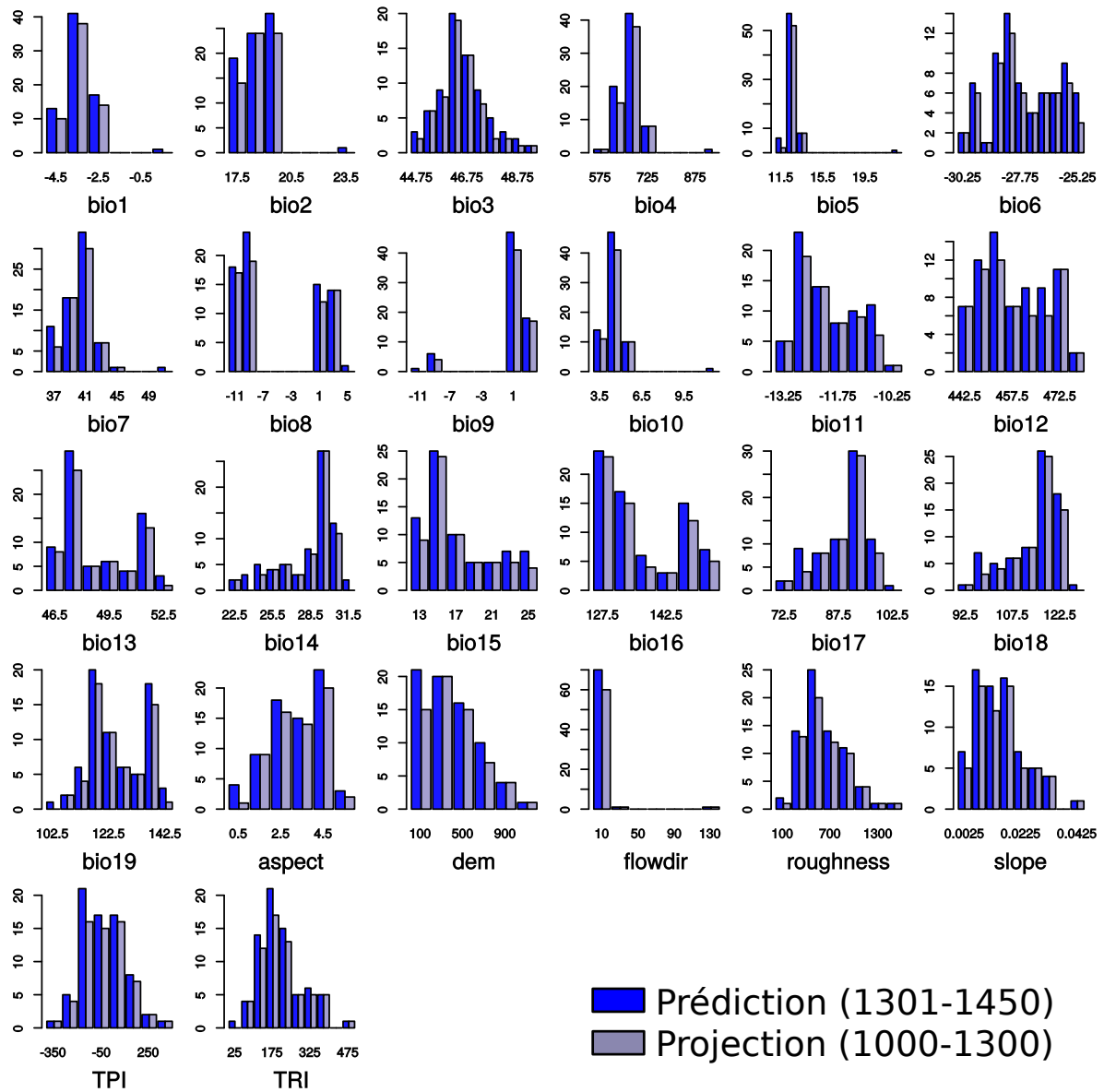


FIGURE 9.14 : Histogrammes des variables bioclimatiques et topographiques du 1^{er} peuplement

FIGURE 9.15 : Histogrammes des variables bioclimatiques et topographiques du 2nd peuplement

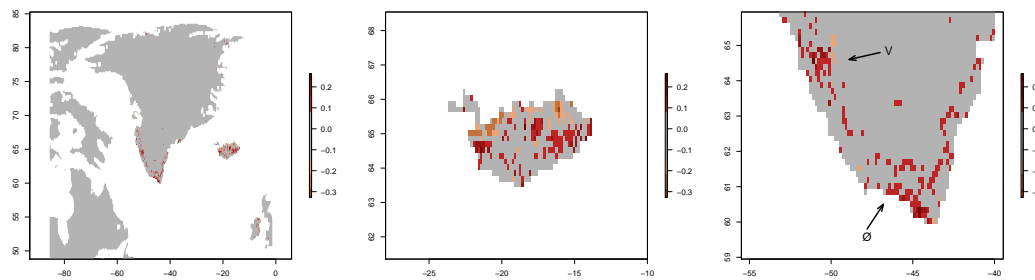
La comparaison d'une prédiction avec une projection de niche nous permet d'évaluer s'il y a eu une conservation de niche. En effet, si l'on observe un changement dans l'expression géographique entre niche prédite et niche projetée en concordance avec les observations concernant la période de projection, nous pouvons dire que la niche est conservée et que les observations de la nouvelle période (les occurrences) sont déterminées par les conditions environnementales. Nous sommes ici en présence du cas contraire où les occurrences de présence disparaissent d'une zone maintenue dans les niches potentielles prédites par l'ECNM, remettant en cause la conservation de niche. Notons tout de même que ces nuances, peu nombreuses, sont localisées en majorité à proximité des colonies (voir figure 9.16).

Entre la prédiction et la projection de la niche du 1^{er} peuplement, on observe majoritairement une diminution du nombre et des valeurs de présence prédites. Pour la niche du 1^{er} peuplement, nous observons que les différences les plus prononcées se situent en Islande. Or, la niche prédite en Islande pour le 2nd peuplement est très peu étendue et de très faible probabilité. Ce qui suggère que la modification de niche a principalement affecté les sites islandais. Nous constatons une situation équivalente mais un peu plus atténuée dans la région du *Vesterbygden*. Cette observation met en évidence un phénomène de sélection induit par un changement climatique. On assiste à un goulot d'étranglement conduisant à une contraction de la niche écoculturelle.

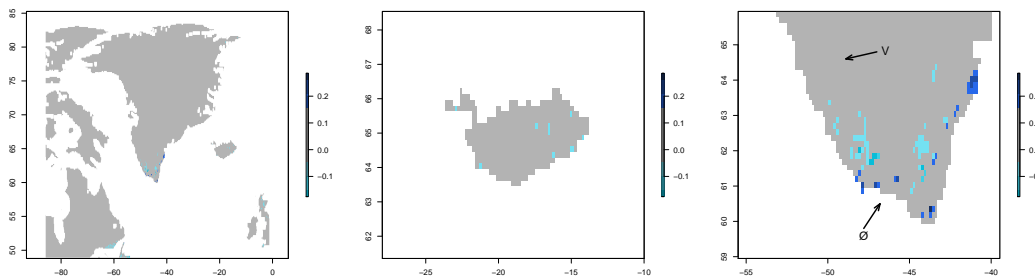
Nous constatons une différence importante entre les niches prédites de chaque peuplement. L'ACP commune (figure 9.11) et les projections géographiques des prédictions des niches (figures 9.12(a) et 9.12(c)) montrent également une sélection de sites et suggèrent une restriction de niche. Les aires des ellipses indiquent que l'espace écologique occupé par la niche du 1^{er} peuplement est beaucoup plus important que celui occupé par celle du 2nd peuplement. Nous voyons aussi que cette niche est entièrement incluse dans la première. De plus, l'observation des histogrammes, figure 9.17, révèle une nette diminution de la largeur de niche (reflétée par l'emprise sur l'axe des abscisses de chacun des histogrammes) en plus d'une diminution d'extension de niche (dont on peut se faire une représentation en cumulant la hauteur des bâtons de chaque histogramme). Ces observations valent pour la plupart des variables bioclimatiques ainsi que pour la variable flowdir (l'écoulement). Les autres variables topographiques ont quant à elles des répartitions sensiblement identiques à celles observées pour le 1^{er} peuplement.

Les comparaisons entre niches prédites et niches projetées dénotent une diminution de l'aire géographique et non de l'espace environnemental car la conservation des caractères environnementaux de la niche est caractéristique de la projection. En revanche, l'expression géographique indique une variation de l'environnement qui est ici le reflet de la détérioration climatique.

Les histogrammes de la figure 9.17 montrent pour plusieurs variables (bio1, bio2, bio3, bio5, bio7, bio12, bio13, bio14, bio15, bio16, bio17, bio18 et bio19) une distribution bimodale des prédictions de présence de la niche du 1^{er} peuplement. Nous constatons qu'un seul des maxima de ces distributions est retenu lorsqu'on observe les distributions



(a) Variations des modèles de niches du 1^{er} peuplement. Delta entre la prédiction (niche établie avec l'environnement de la période 1000-1300 AD) et la projection de cette niche dans l'environnement de la période suivante (1301-1450 AD). Les tons rouges représentent un delta positif (c'est-à-dire des pixels où la présence prédite a une probabilité plus élevée dans la prédiction que dans la projection) ; en dégradé d'orange les lieux où la projection indique une probabilité de présence supérieure à la prédiction.



(b) Variations des modèles de niches du 2nd peuplement. Delta entre la prédiction (niche établie avec l'environnement de la période 1301-1450 AD) et la projection de cette niche dans l'environnement de la période précédente (1000-1300 AD). Les tons bleus foncés représentent un delta positif (c'est-à-dire des pixels où la présence prédite a une probabilité plus élevée dans la prédiction que dans la projection) ; en dégradé de bleu clair les lieux où la projection indique une probabilité de présence supérieure à la prédiction.

FIGURE 9.16 : Variations entre niches prédites et projections.

des prédictions de présence de la niche du 2nd peuplement selon les mêmes variables. Ceci met en évidence le phénomène de contraction de niche. Ces différences de distributions permettent de voir quelles sont les variables environnementales les plus influentes lors de la deuxième période.

L'ACP commune (figure 9.11) indique que cette restriction de niche est accompagnée d'un décalage des optima des niches selon les deux axes et en relation avec les précipitations. Afin de voir si les variations de ces facteurs environnementaux ont été homogènes selon les régions, nous avons tracé des courbes de précipitations et de températures (figure 9.18) en quatre points de grille de la simulation SGI correspondant aux colonies vikings d'Islande, de l'Ouest et de l'Est du Groenland ainsi que du Labrador. Nous observons des nuances en fonction des régions au cours du temps pouvant peut-être expliquer les abandons de sites.

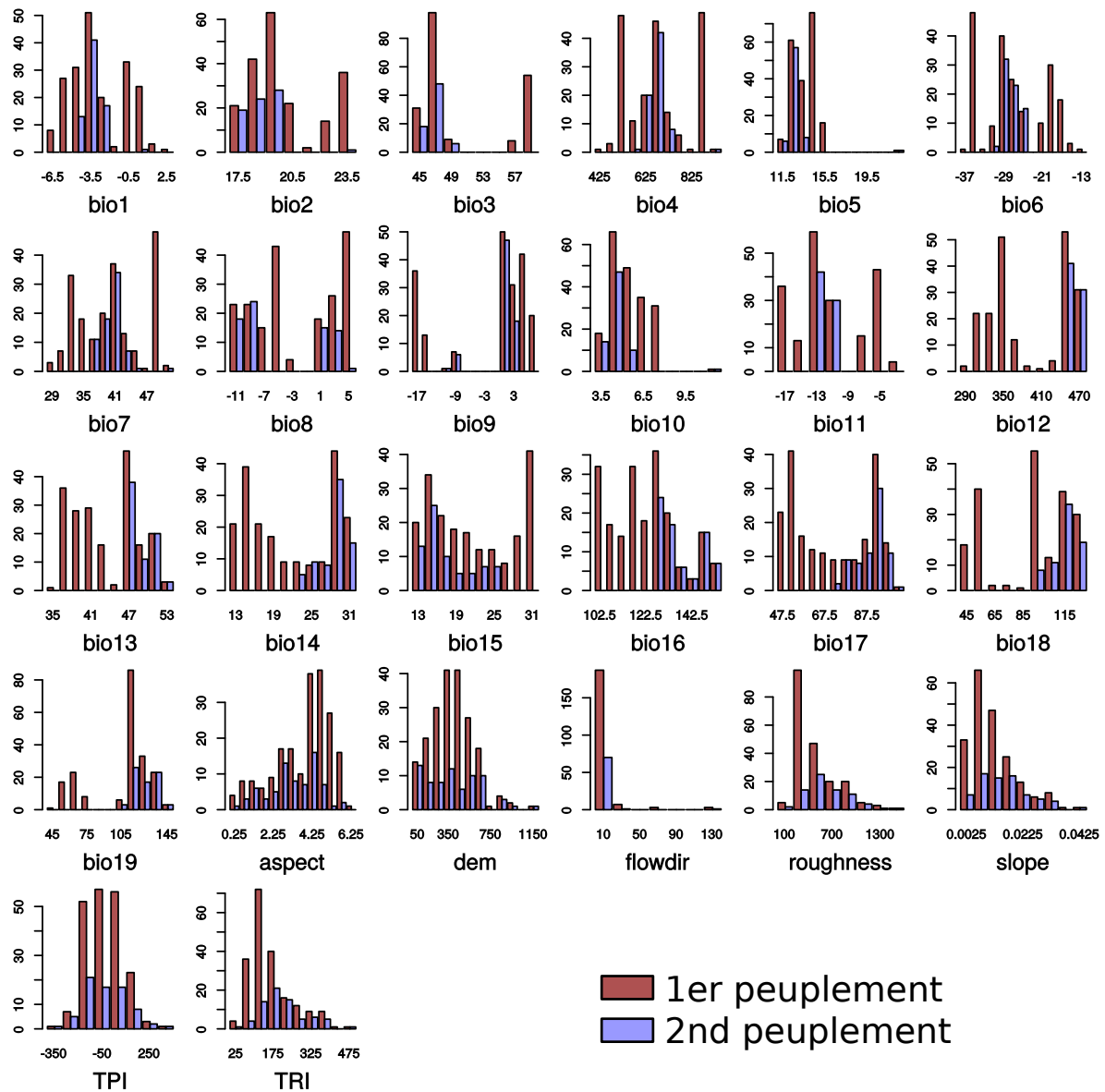


FIGURE 9.17 : Histogrammes des variables bioclimatiques et topographiques des 1^{er} et 2nd peuplements

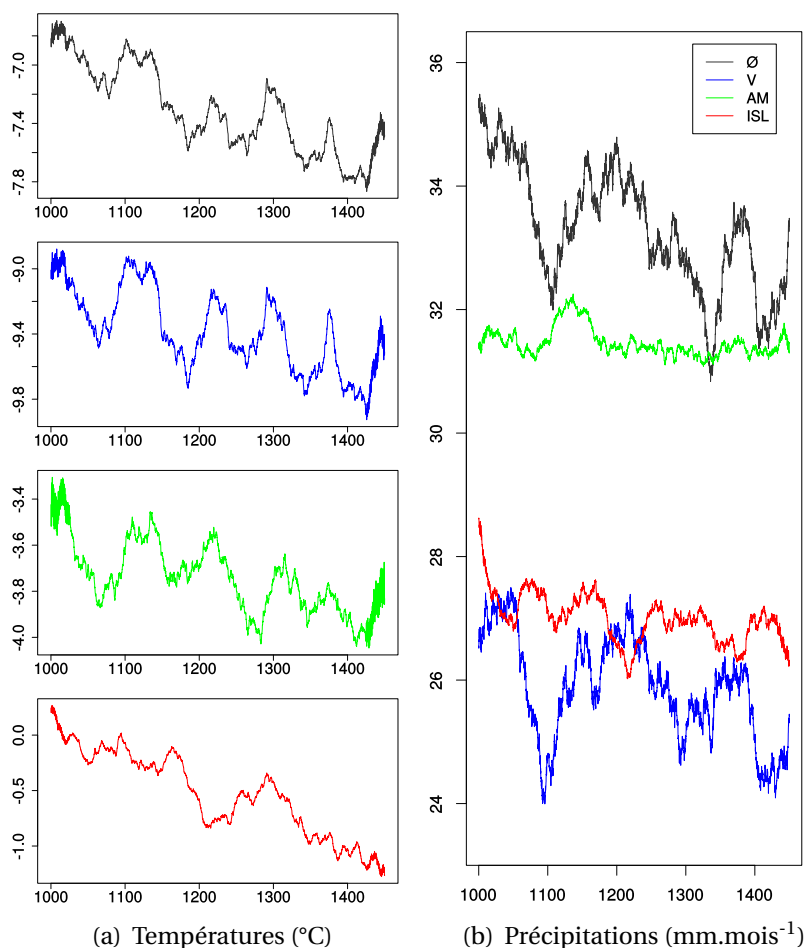


FIGURE 9.18 : Variations des températures et des précipitations locales pour les colonies nordiques du XI^e au XV^e siècle. Les courbes de températures ont été tracées selon les valeurs de la simulation SGI [Servonnat, 2010]. Pour chaque colonie scandinave étudiée : l'Østerbygden (Ø), le Vesterbygden (V), l'Anse-aux-Meadows (AM) et l'Islande (ISL). Du fait de leur proximité géographique et de la résolution de la simulation SGI, les courbes V représentent aussi les variations de températures et précipitations du *Mellemygden* au cours de ces siècles.

Afin de déterminer quelles sont les variations des niches écoculturelles dans le temps, nous avons reconstruit et projeté des modélisations selon un pas de temps séculaire.

9.3.3.6 Dynamique séculaire et variabilité locale des niches écoculturelles vikings

Projections séculaires

La période la plus favorable à la colonisation viking de ces régions est le XI^e siècle. Nous reconstruisons alors la niche écoculturelle pour cet environnement en utilisant les occurrences qui correspondent à leur maximum d'extension, c'est-à-dire celles du 1^{er} peuplement. Puis nous la projetons dans les siècles suivants (XII^e à XV^e siècles) afin d'apprécier plus finement l'évolution de cette niche. De même, nous projetons la niche du 2nd peuplement modélisée pour le XIV^e siècle et dans ceux qui précèdent (XI^e à XIII^e siècles) ainsi qu'au XV^e siècle. Les niches écoculturelles vikings correspondantes sont représentées dans l'ACP figure 9.19.

Comme observé précédemment avec le découpage temporel 1000-1300 AD, 1301-1450 AD. Nous remarquons que les variations entre niches prédites et projetées sont subtiles. Cependant, nous notons des variations antagonistes selon les axes du plan de l'ACP. En effet, si l'on observe les reconstructions de niches pour le 1^{er} peuplement, une réduction de l'espace environnemental disponible est suggérée selon l'axe 1 et une augmentation selon l'axe 2. Notons que la restriction est de plus grande ampleur que l'augmentation. Ainsi selon le deuxième axe, en relation avec les températures saisonnières et les précipitations des périodes les plus humides de l'année, les projections de la niche du XI^e siècle indiquent une plus grande surface d'environnement favorable pour toutes les autres périodes, avec un maximum lors du XIII^e siècle. En revanche, selon l'axe 1, qui est le plus représentatif (34.1% de la variance expliquée), toutes les projections signalent une restriction de la surface convenable à la niche prédite pour le XI^e siècle. L'expression environnementale de niche la plus réduite est celle correspondant à la projection du XIV^e siècle. L'environnement de cette période serait donc le moins propice à la culture viking du XI^e siècle.

La projection de la niche du XI^e siècle au XII^e siècle est relativement proche de celle du XI^e siècle ; ce résultat montre que l'environnement n'a pas changé entre les deux siècles. Nous notons toutefois une restriction selon l'axe 1 que l'on retrouve dans les XIII^e et XV^e siècles. Pour la projection de la niche au XIV^e siècle, on observe une très forte restriction selon l'axe 1. Cette réduction de niche est en accord avec les faits historiques (abandon du *Vesterbygden*). Pour expliquer le fait que la colonie du *Vesterbygden* fut abandonnée alors que les conditions environnementales permettaient à nouveau son exploitation, nous émettons l'hypothèse d'un décalage de niches s'amorçant dès le XII^e siècle. En effet, grâce à une certaine plasticité de leur niche écoculturelle, les colonies ont une capacité de résilience, mais ce phénomène serait parallèlement accompagné d'une adaptation induite par les changements environnementaux. Ce décalage de niches permet de compenser la constriction environnementale observée sur le premier axe de l'ACP pour les XII^e et XIII^e siècles.

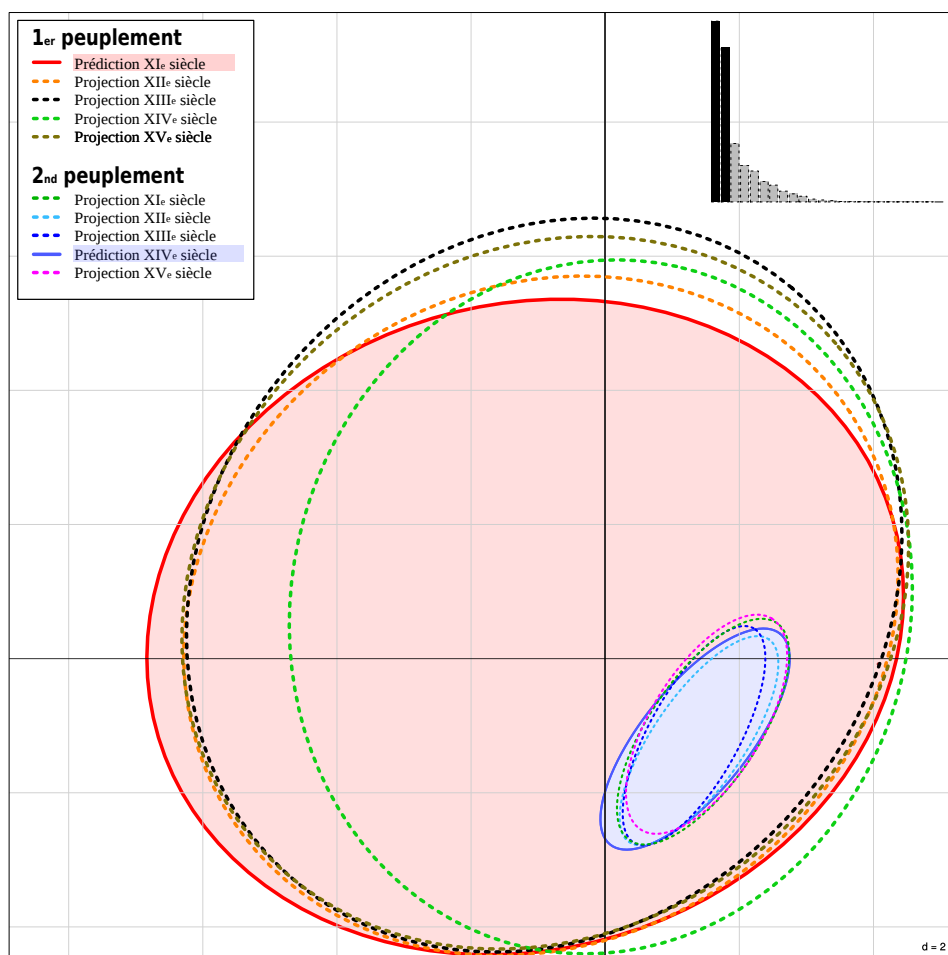


FIGURE 9.19 : Dynamique environnementale séculaire des niches des 1^{er} et 2nd peuplements. Les ellipses sont construites à partir des prédictions de présence ayant une probabilité supérieure à 0.9. La niche du 1^{er} peuplement a été prédite avec l'environnement du XI^e siècle et projetée dans les siècles suivants alors que la niche du 2nd peuplement a été prédite pour le XIV^e siècle puis projetée dans les autres siècles. Le premier axe représente 34.1% de la variance expliquée et le deuxième axe 29.11%. L'histogramme des valeurs propres est représenté dans le coin supérieur droit.

Cependant, cette dernière est trop élevée au XIV^e siècle et provoque le déclin irréversible des colonies vikings qui commence par l'abandon du *Vesterbygden*. Ainsi, le système adaptatif viking des régions nordiques aurait amorcé une mutation qui ne lui permet pas une résilience au XV^e siècle, période pendant laquelle l'espace environnemental convenable à la niche écoculturelle optimale viking était aussi large qu'au XIII^e siècle.

Nous estimons également que le système viking des régions nordiques reposait sur un équilibre fragile basé sur les échanges entre ses différentes régions. Or, nous voyons figure 9.18 que les conditions en Islande n'ont cessé de continuellement se dégrader tout au long de ces périodes d'occupation viking, déstabilisant d'autant plus ce système. L'*Østerbygden*, ne bénéficie alors plus des apports ni du *Vesterbygden*, abandonné, ni de l'Islande, où les conditions se dégradent. Il périclite à son tour, probablement autour du XV^e siècle.

Variabilité locale

La figure 9.18 illustre les variations de températures et de précipitations spécifiques aux colonies au cours des cinq siècles d'occupation viking des zones nordiques. Les variations de températures sont marquées entre l'Islande et les autres contrées. Les températures sont quant à elles différentes pour chaque région. Le déclin de la température est progressif pour les quatre régions, néanmoins les courbes du Groenland et de l'Anse-aux-Meadows présentent des oscillations de périodes chaudes et de périodes froides. Nous suggérons que ces périodes chaudes ont constitué des phases pendant lesquelles les colonies du Groenland ont pu maintenir leur système de production à l'inverse de l'Islande qui, comme nous pouvons le constater sur la figure 9.18, subit une détérioration climatique quasi-continue du XI^e au XV^e siècle. L'Islande connaît vers 1200 AD une chute importante de températures. Cet épisode froid n'a pas lieu au même instant à l'Anse-aux-Meadows. Les vagues de froid pouvant couper la route entre l'Islande et le Groenland [Kuijpers *et al.*, 1999], nous suggérons qu'un approvisionnement en ressource depuis le Labrador a pu avoir lieu pendant cette période pour pallier les éventuelles interruptions d'échanges avec l'Islande.

Nous avons vu précédemment que les précipitations sont probablement le facteur déclencheur du décalage de niches. Ces dernières sont plus importantes pour l'Anse-aux-Meadows et pour l'*Østerbygden* que pour l'Islande et le *Vesterbygden* sur toute la période. Nous notons leur stabilité à Anse-aux-Meadows avec cependant un léger pic autour de 1100 AD. En Islande les précipitations chutent à partir de 1150 AD et atteignent leur minimum en 1250 AD, date à laquelle on observe une moyenne annuelle de 26 mm/mois (figure 9.18). La sécheresse précoce de l'Islande aurait pu entraîner son déclin et déséquilibrer le système. Les colonies du Groenland endurent à leur tour, près de 50 ans plus tard, un épisode sec avec un minimum des précipitations autour de 1300 AD, date du début de l'abandon de la colonie de l'ouest. Seules les colonies de l'*Østerbygden* et Anse-aux-Meadows avec de plus fortes précipitations se maintiennent grâce à une rehausse des précipitations mais une dernière sécheresse en 1400 AD aurait pu leur être fatale.

Lorsqu'on représente les différences géographiquement entre niches prédites et niches projetées (figure 9.16) nous nous apercevons que les variations sont plus marquées pour les modélisations concernant le 1^{er} peuplement. Ces variations sont principalement marquées pour les colonies islandaise et groenlandaises. Le delta, différence entre les valeurs des prédictions et les valeurs des projections, est largement positif. Il indique une diminution de la capacité d'accueil de l'environnement pour la niche écoculturelle du 1^{er} peuplement. Les variations entre les niches prédites et projetées sont moins marquées en ce qui concerne les modélisations du 2nd peuplement. Le delta est cette fois majoritairement négatif, indiquant également un environnement moins convenable à la niche viking. Les faits historiques nous indiquent que seul l'*Østerbygden* résiste (au moins un siècle) à la détérioration climatique, et qu'il a existé un possible refuge au Labrador. L'ECNM détecte aussi cette possibilité bien qu'en ce lieu les prédictions de présence soient de faibles probabilités figure 9.9.

Ce système tripolaire permet une niche large et viable dans ces régions, dépendante du climat. Cependant, cette niche écoculturelle ne peut supporter que des variations environnementales de faible amplitude. Or, nous observons des variations abruptes concomitantes à plusieurs régions et suggérons qu'elles sont à l'origine d'un déséquilibre qui a mis en péril le maintien de la niche. L'*Østerbygden* en position centrale bénéficie d'une place privilégiée en termes d'échanges économiques. Cette localisation a pu jouer un rôle dans la résilience de la culture viking et dans sa résistance tardive comparée à celles des autres colonies. Ce phénomène de « résilience/résistance » pourrait s'expliquer par les conditions climatiques plus douces de la colonie de l'Est. La généralisation de la détérioration limite la résilience de cette colonie qui se maintient quelques décennies puis disparaît. Un refuge en Amérique n'offrirait pas d'après l'ECNM des conditions suffisamment favorables à la résilience du système. De plus des facteurs culturels entrent en jeu. En effet, les interactions avec les *screaling* et avec les Inuits auraient pu faire obstacle aux possibilités de refuges.

9.4 Interactions culturelles

Vers le XIV^e siècle, une autre culture, provenant d'Amérique du Nord vient s'établir au Groenland, il s'agit des Inuits. L'expansion de ces derniers, bien plus adaptés aux conditions extrêmes, a certainement été favorisée par la détérioration climatique du PAG. Les Vikings et les Inuits ont pu être, un temps donné, en compétition pour les zones dont les caractéristiques environnementales pouvaient correspondre à chacune de leurs niches (chevauchement de niches). Bien que le contact des deux sociétés soit indéniable [Golding *et al.*, 2011; Schledermann et Mc Cullough, 2003] aucune preuve archéologique, à notre connaissance, ne montre une hybridation de culture. Cependant des études des restes fauniques et isotopiques menées sur des ossements humains [Arneborg *et al.*, 2012a,b] nous apprennent, qu'à partir du XIII^e siècle, les Vikings se nourrissaient plus de produits de la mer que de la terre.

Reconstruire les niches écoculturelles des Inuits et comparer leurs évolutions avec celles

des niches vikings permettrait d'estimer la part de leurs interactions dans la disparition des Vikings. L'intégration à l'étude de groupes humains contemporains vivant à proximité de l'Anse-aux-Meadows pourrait expliquer pourquoi cette région n'a pas pu servir de refuge à la culture Viking.

***Synthèse du chapitre : Impact du changement climatique sur les
populations scandinaves du Groenland***

La modélisation des niches écoculturelles des Vikings des régions subarctiques permet de confirmer que leur disparition est principalement due à l'avènement du Petit Âge Glaciaire. Elle met également en évidence un goulot d'étranglement qui provoque la disparition des colonies les unes après les autres.

L'analyse des ECN séculaires offre une vision détaillée de l'évolution de la niche viking au cours du temps. Comparée aux variations de températures et de précipitations, elle évoque une capacité de résilience du mode de vie viking rendue possible par un système tricéphale. Ainsi, l'observation fine des variations climatiques locales et nos connaissances sur les interactions entre sites d'une même culture, mais aussi interculturelles nous permettent de discuter la part de l'environnement et des relations sociales dans les capacités de résilience et de résistance d'une niche écoculturelle face à un changement climatique.

Nous avons vu dans ce chapitre que l'ECNM des Vikings fournissait des résultats en accord avec les faits historiques faisant de cette étude un témoin validant l'emploi de l'ECNM pour des populations du passé et montrant que l'impact du climat sur une population peut être évalué par la modélisation de niche.

Modélisation des niches écoculturelles des cultures initiales du Néolithique

Préambule

L'adoption du mode de vie néolithique représente la plus grande rupture économique qu'ont connue nos sociétés. Le passage d'une subsistance basée sur la prédation et sur la collecte à une économie de production a été étudié sous bien des coutures : démographie, compétition, inégalités sociales ou encore changements climatiques. Ce chapitre porte sur l'étude des relations de trois grandes cultures du Néolithique ancien avec l'environnement et sur la conservation de leurs systèmes adaptatifs dans les cultures du Néolithique moyen. Il traite dans un premier temps de la capacité de l'ECNM à mettre en évidence des facteurs environnementaux favorisant la propagation de la néolithisation ; puis dans un second temps de sa faculté éventuelle à fournir des arguments pour discuter des héritages culturels.

Sommaire

10.1 Les grands courants de néolithisation en Europe	225
10.2 Niches écoculturelles des cultures du Néolithique ancien	228
10.3 Nouvelle étude	232
10.4 Perspectives apportées par la nouvelle étude	243

10.1 Les grands courants de néolithisation en Europe

La plupart des auteurs s'accordent sur l'apparition du phénomène de néolithisation au Proche-Orient il y a environ 11500 ans. Cette transition procure des changements d'ordre démographique [Bocquet-Appel, 2008; Fort et Méndez, 1999]), idéologique [Banning et Byrd,

1987; Cauvin, 1994; Descola, 2004] et techno-économique (développement de l'agriculture et de l'élevage) [Arias Cabal, 2007; Bonsall *et al.*, 1997; Guilaine, 2003; Marchand et Manen, 2006; Zvelebil, 2000].

Plusieurs hypothèses mises en exergue pour expliquer le phénomène de transition du Mésolithique au Néolithique sont fortement débattues. Des hypothèses dites « migrationnistes », expliquant la mise en place du peuplement européen en plusieurs vagues colonisatrices [Ammerman et Cavalli-Sforza, 1971] mettent en avant le déploiement de l'économie agro-pastorale par la colonisation d'individus provenant du Levant. Ces dernières sont opposées aux hypothèses dites « culturalistes » qui promeuvent l'acculturation des derniers Mésolithiques [Zvelebil, 2000], certains auteurs avancent aussi une découverte autochtone de l'agriculture dans certaines régions [Henrard, 2003; Marchand, 2003; Otte et Noiret, 2001].

La migration d'individus est la première explication fournie à la diffusion de ce nouveau mode de subsistance [Ammerman et Cavalli-Sforza, 1971; Childe, 1925; Clark, 1965]. Des arrivées de groupes néolithiques originaires du Proche-Orient sont attestées par des études portant sur des caractères biologiques, (par exemple d'ordres craniométrique [Pinhasi, 2006] ou génétique [Ammerman *et al.*, 2006; Chikhi *et al.*, 1998, 2002; Cymbron *et al.*, 2005; Haak *et al.*, 2010, 2005]). La progression de la néolithisation est également mise en évidence par la culture de plantes non endémiques et l'élevage d'espèces animales allochtones, originaires du Proche-Orient, retrouvées dans des sites d'Europe de l'Ouest [Götherström *et al.*, 2005; Kreuz *et al.*, 2005; Rösch *et al.*, 2002; Vigne, 1993, 1998, 2000, 2008; Zapata *et al.*, 2004; Zeder, 2005, 2006, 2008]. Toutefois, les données biologiques révèlent également des indices de la participation d'espèces animales autochtones au développement de l'élevage en Europe [Larson *et al.*, 2007; Randi, 2008].

Plus d'un siècle après la définition du terme Néolithique [Lubbock, 1865], la transition d'une économie basée sur la prédation, la cueillette et la collecte, vers une économie de production (d'agriculture et d'élevage), est perçue comme complexe, asynchrone, multi-directionnelle et vérifiant à la fois des théories migrationnistes comme des théories culturalistes. Ce débat est alimenté également par des approches linguistiques qui ne font pas consensus [Gray et Atkinson, 2003; Renfrew, 1990, 1996].

Ce nouveau mode de subsistance est accompagné de la domestication d'espèces végétales et animales. Les agriculteurs de cette période cultivent des espèces d'orge, de blé, de lentilles, et de pois dans des régions où le climat et la saisonnalité sont différents de leurs lieux d'origines [Kreuz *et al.*, 2005]. Ce dernier point requiert par conséquent une parfaite connaissance des contraintes environnementales qu'impose la culture des espèces domestiquées et une recherche de terres sous des climats propices à leurs cultures. L'élevage de bovidés et de suidés implique le même savoir. La diffusion de ces sociétés peut donc en partie être liée aux conditions environnementales [Mohen, 2006], ou au fait que ces groupes humains ont une maîtrise des espèces domestiquées suffisamment importante au point de pouvoir s'affranchir des exigences biologiques de ces dernières. Par ailleurs, leur forte capacité d'adaptation à de nouvelles contraintes environnementales est marquée, à plusieurs

reprises, par l'intégration d'espèces endémiques à leurs cheptels (introduction résultant de domestications ou de croisements inter-espèces). Ainsi, la néolithisation advient comme une véritable révolution technique, économique et culturelle [Barker, 2006; Cauvin, 1994; North et Thomas, 1977] qui se diffuse en Anatolie puis à travers l'Ancien Monde [Kertész et Sümegei, 2001; Kreuz, 2007; Kreuz *et al.*, 2005]. L'adoption d'un système de production sur l'ensemble d'un continent dont l'environnement est hautement diversifié comme l'Europe, relève d'un succès sans pareil. Cet exploit soulève le questionnement des divers contextes environnementaux affrontés par ces sociétés ainsi que des stratégies adaptatives mises en place pour aboutir à la large diffusion du Néolithique.

Deux voies de néolithisation sont observées : une maritime le long de la côte méditerranéenne (qui comprend les cultures Impressa et Cardial) et une continentale suivant le Danube associée à la culture LBK (*Linerbandkeramik*) [Cymbbron *et al.*, 2005; Paschou *et al.*, 2014; Perlès, 2010; Price, 2000; Whittle, 1996] (voir figure 10.1).

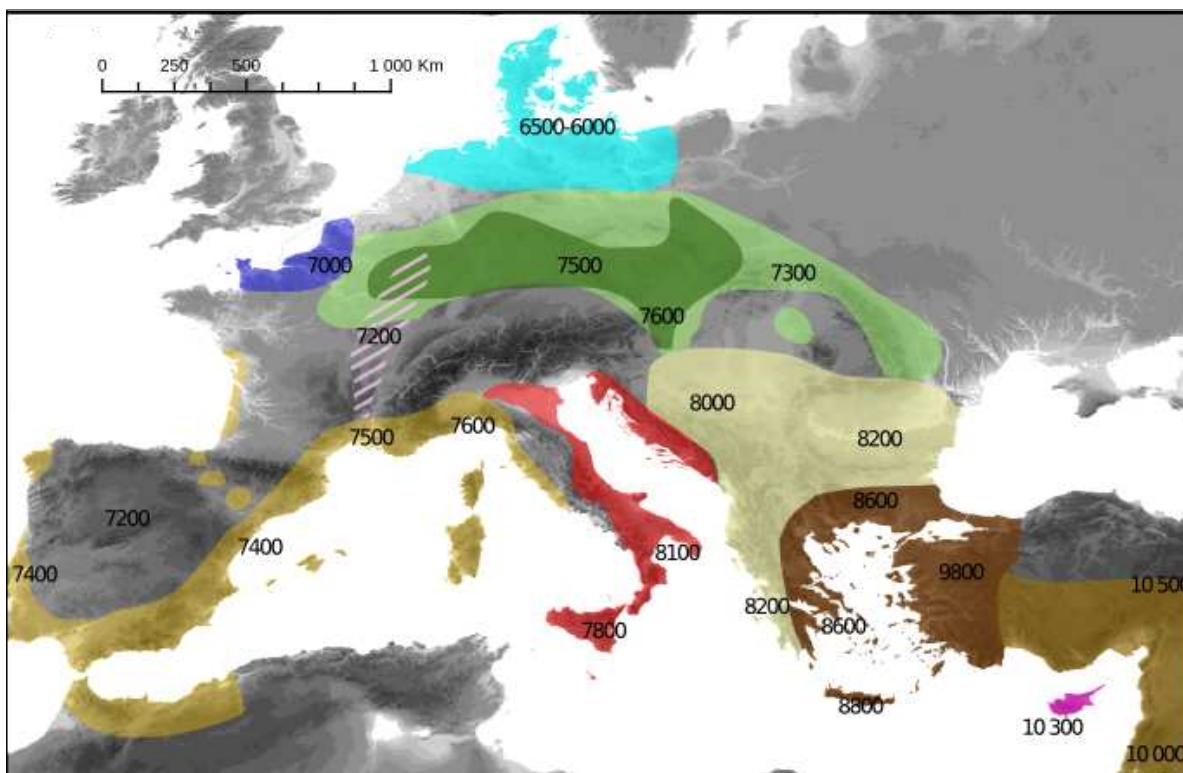


FIGURE 10.1 : Cultures archéologiques majeures identifiées au Néolithique ancien en Europe. En rouge : aire d'extension de l'Impressa, en jaune : aire du Cardial, en vert : extension du rubané (les nuances de vert reflètent deux étapes d'expansion successives). D'après [Guilaine, 2003] (modifié, DAO S. Rigaud) ; dates indiquées en cal BP

Ces deux courants culturels se distinguent notamment par leurs céramiques, leurs activités agricoles, leurs industries lithiques ainsi que par leurs architectures du bâti [Burger et Thomas, 2011; Coudart, 1993; Dolukhanov *et al.*, 2005]. Ces distinctions laissent supposer que leur économie, voire leurs modes de vie en général, doivent être également différents.

La néolithisation de l'Europe est majoritairement accomplie au Néolithique moyen (il y a environ 6000 BP). On assiste dès lors à un buissonnement de cultures à l'échelle du continent. Ces dernières se démarquent par la diversification de production matérielle, de pratiques funéraires et le déploiement de vastes réseaux d'échanges complexes et diversifiés [Cauwe *et al.*, 2007; Jeunesse, 2004; Sidera, 1997, 2000; Spatz, 1997].

Bien que l'origine et les modes de transmission culturelle de l'Europe du Néolithique ancien aient été fort débattus, peu d'études ont porté sur la question de la relation spécifique entre ces grands courants culturels et l'environnement. Dans l'objectif de savoir à quels challenges environnementaux ont fait face les populations en voie de néolithisation et dans quelles mesures ces derniers ont influencé leur diffusion puis leur diversification, nous avons reconstruit les niches écoculturelles des cultures méditerranéennes et danubiennes.

10.2 Niches écoculturelles des cultures du Néolithique ancien

Le travail présenté dans ce chapitre s'appuie sur une étude antérieure [Antunes, 2010; Banks *et al.*, 2013a]. Nous complétons cette étude en modélisant les ECN des différentes cultures du Néolithique grâce à 12 algorithmes prédictifs et un jeu de variables environnementales plus conséquent et plus précis. Avant d'aborder les nouvelles modélisations, qui permettent de soulever de nouvelles questions (section 10.3.2), il est nécessaire d'avoir connaissance des éléments majeurs apportés par la première étude. Cette section en présente donc une synthèse.

10.2.1 Matériels et méthodes

10.2.1.1 Occurrences

Nous avons sélectionné un corpus réunissant des sites archéologiques affiliés aux cultures de l'Impressa, du Cardial et du Danubien (LBK) [Antunes, 2010; Banks *et al.*, 2013a]. L'attribution culturelle de chacun des sites a été validée en fonction de leurs dates et du matériel archéologique exhumé. Dans le but d'obtenir une représentativité culturelle optimale, les sites dont l'attribution culturelle était discutable ont été exclus de l'analyse. De cette manière, 90 sites ont été retenus pour le LBK, 22 sites pour l'Impressa et 32 sites pour le Cardial (figure 10.2); voir suppléments de l'article pour les références associées aux sites. Ces sites couvrent l'Europe, de la Hongrie à la France pour la culture LBK (ou danubienne), et le Sud

de l'Europe, de l'Italie au Portugal pour les cultures méditerranéennes de l'Impressa et du Cardial.

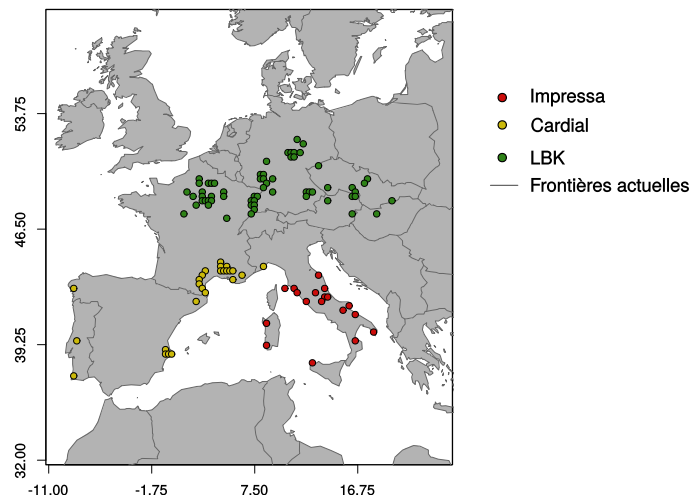


FIGURE 10.2 : Occurrences des sites néolithiques utilisés pour l'ECNM.

10.2.1.2 Définition d'une zone M

Un M a été appliqué pour optimiser les prédictions (voir section 2.4.2.2). Dans un premier temps pour les cultures du Cardial et de l'Impressa, le M recouvrait l'Europe entre les latitudes 30° et 50°N environ. De cette manière, en réduisant l'aire de la modélisation, nous nous assurons de pouvoir déceler au mieux les éventuelles différences entre les niches écoculturelles de ces deux faciès culturels. Pour le deuxième volet de l'étude, qui consistait à comparer la niche méditerranéenne à la niche danubienne, le M utilisé couvrait l'Europe entre les latitudes 30° et 60°N environ, englobant cette fois des régions plus septentrionales comme la Grande-Bretagne (figure 10.3(b)).

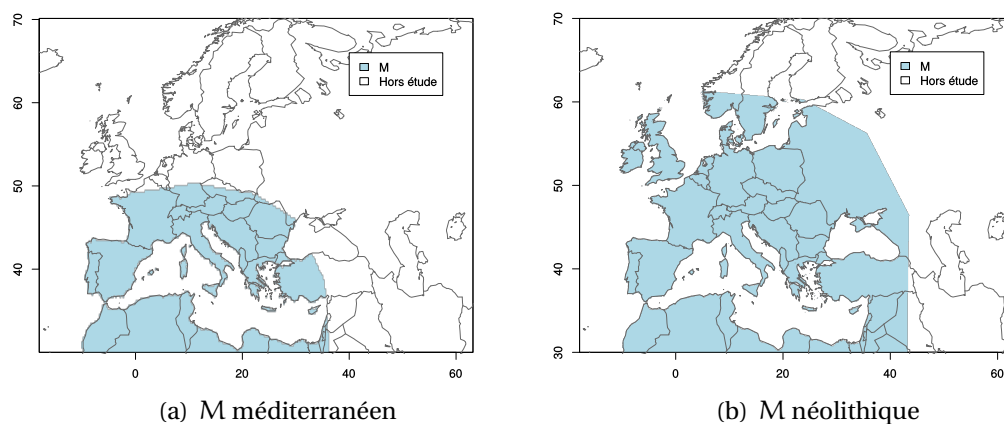


FIGURE 10.3 : Cartes des M utilisés pour les ECNM du Néolithique. Les ECNM de [Banks *et al.*, 2013a] utilisent les deux M, alors que l'autre étude n'utilise que le M néolithique (b).

10.2.1.3 Données environnementales

Les données climatiques sont issues d'une simulation paléoclimatique pour l'Holocène moyen (6000 BP) [Sepulchre *et al.*, 2008]. Ce modèle zoomé (LMDZ4, [Li et Conil, 2003]) est focalisé sur le Tchad avec une résolution de 50 km. Ce qui confère pour l'Europe une résolution variant de 70 à 90 km de côté. Aux enveloppes climatiques issues du modèle, ont été couplées des variables topographiques provenant du jeu de données Hydro-1K (U.S. Geological Survey, Center for Earth Resources Observation and Science, <http://eros.usgs.gov/>) et altimétriques depuis le jeu de données ETOPO1 [Amante et Eakins, 2009]. Les variables environnementales issues de ce modèle et utilisées pour l'ECNM sont les suivantes :

- la température annuelle moyenne
- les précipitations annuelles moyennes
- la température du mois le plus chaud
- la température du mois le plus froid
- l'altitude
- la pente
- l'index topographique
- et l'orientation de la pente.

Les données environnementales et d'occurrences archéologiques ont été ensuite injectées dans les algorithmes prédictifs de niches écologiques.

10.2.1.4 Algorithmes prédictifs utilisés

L'ECNM a été réalisée grâce aux deux algorithmes les plus fréquemment employés en modélisation de niche, à savoir DesktopGARP et MaxEnt (voir section 2.2.5).

Les ECN obtenues ont été ensuite analysées grâce à des analyses en composantes principales (ACP) et comparées entre elles grâce au logiciel ENMTools [Warren *et al.*, 2010].

10.2.2 Principaux résultats et discussion

Malgré les tendances « microprédictive » de MaxEnt et « macroprédictive » de GARP (voir sections 2.2.5.1 et 2.2.5.2) les niches écoculturelles obtenues avec les deux algorithmes restent comparables pour l'ensemble des résultats.

Il ressort de cette étude que les niches écoculturelles de l'Impressa et du Cardial sont très similaires. Les deux cultures peuvent être alors considérées comme ayant la même écologie, le même système adaptatif. L'hypothèse de contraintes environnementales différentes pour expliquer les localisations géographiques différentes des deux cultures peut être rejetée. À la lumière de ce résultat, les deux niches ont été combinées en une seule reflétant l'adaptation méditerranéenne.

Les niches des Danubiens et des Méditerranéens apparaissent dans ce travail comme étant radicalement distinctes l'une de l'autre hormis en de rares lieux qui représentent au total une surface très réduite. On note une niche plus large pour les Méditerranéens que pour les Danubiens, ces résultats signifient que les contraintes écologiques sont perçues différemment par chacun des courants de néolithisation. En effet les ACP montrent que les températures et les précipitations sont les paramètres qui jouent les plus grands rôles dans la composition des niches, avec cependant des relations opposées entre ces facteurs environnementaux pour chaque entité culturelle. L'isotherme 8.5°C de température moyenne annuelle apparaît comme longeant la frontière entre les niches des deux grands courants de néolithisation étudiés. La niche du LBK est caractérisée par des précipitations annuelles variant de 700 à 1800 mm/an contre 360 à 1100 mm/an pour la niche des Méditerranéens. L'étude a dévoilé également que les variables topographiques jouent un rôle secondaire mais néanmoins important. L'expression géographique de la niche de la culture danubienne comprend des lieux généralement situés sur des paysages réguliers à une altitude inférieure à 500 m alors que la niche des Méditerranéens a une distribution qui inclut de larges variations topographiques.

Montrant une concordance des niches écoculturelles avec les aires de distribution des différentes cultures établies par les préhistoriens, cette étude met en évidence que l'établissement de ces groupes humains du Néolithique est en grande partie lié à l'environnement.

Plusieurs études ont montré que ces sociétés cultivaient des espèces adaptées à leur environnement [Bakels, 2014; Kreuz, 2007]. Ainsi, les contrastes environnementaux entre les aires de répartition des niches écologiques de ces deux courants de néolithisation ont certainement induit des agricultures variées. Ce travail montre également que la diversité

des espèces cultivées est en relation avec la diversité environnementale et donc la largeur de niche, expliquant ainsi une plus grande variété agricole dans les cultures méditerranéennes [Zapata *et al.*, 2004].

Enfin, nous proposons l'existence de raisons culturelles, technologiques ou démographiques pouvant restreindre l'expansion des cultures dans la totalité de leurs niches prédites par les algorithmes. Sont évoqués par exemple le fait que les Danubiens n'atteignent pas les îles britanniques ou que le courant Méditerranéen ne se développe pas le long de la côte méditerranéenne d'Afrique du Nord. Notons que certains auteurs attribuent des sites archéologiques au Cardial en Afrique du Nord [Manen *et al.*, 2007; Sánchez *et al.*, 2012; Zilhão, 2014], néanmoins leur quantité reste faible comparé à la surface que les Méditerranéens pouvaient occuper d'après la prédiction de leur niche et bien que ces groupes humains possédaient la technologie de navigation leur permettant d'y accéder [Ammerman, 2011; Zilhão, 2014].

Au fil de cette thèse nous avons élaboré des techniques afin d'automatiser la modélisation et de combiner les résultats des algorithmes prédictifs (chapitre 2). Nous avons remarqué que les variables bioclimatiques pouvaient s'avérer performantes pour reconstruire les niches écoculturelles (chapitres 8 et 9). Dans le chapitre 1, nous avons noté que la résolution des variables environnementales pouvait fortement influencer les résultats d'une modélisation. Forts de cela, nous avons décidé de reconduire notre étude [Antunes, 2010; Banks *et al.*, 2013a] en incorporant ces acquis méthodologiques.

10.3 Nouvelle étude : données complémentaires et approfondissement

Nous avons décidé pour la nouvelle étude de conserver les sites archéologiques et les attributions culturelles de la première (voir section 10.2.1.1 et figure 10.2). En revanche, nous avons augmenté le nombre de variables environnementales ainsi que leur résolution spatiale. Nous avons également utilisé plusieurs techniques de modélisation et réalisé des consensus élitistes pondérés (voir chapitre 6).

Dans la première étude, la comparaison des cultures Impressa et Cardial nécessitait un M adapté afin de pouvoir les comparer : le M méditerranéen (figure 10.3(a)). Dans la nouvelle étude, nous cherchons à connaître leurs positions écologiques relatives par rapport aux Danubiens. Nous avons, pour toutes les modélisations, utilisé le même M : le M néolithique (figure 10.3(b)).

10.3.1 Matériels et méthodes de la nouvelle étude

10.3.1.1 Ajout de variables environnementales et augmentation de la résolution spatiale

La première étude nous a permis de constater que les grandes cultures du Néolithique ancien étaient influencées par les températures, par les précipitations et dans une moindre mesure par la topographie. Les stratégies de subsistance des groupes humains étudiés sont principalement basées sur l'agriculture et l'élevage. Ces sociétés doivent donc gérer leur production vivrière au gré de ces variables qui, dans les latitudes concernées, évoluent au fil de l'année. Dans le but de mieux cerner les relations cultures climats l'utilisation de variables bioclimatiques (tableau 2.6) s'avèrent donc plus adaptée que les températures et les précipitations annuelles car elles prennent en compte les variations saisonnières de ces deux facteurs.

Les données bioclimatiques ont été calculées à partir de la simulation climatique *WorldClim* à 6000 BP. Cette simulation est issue d'un *downscaling* du modèle CCSM4 [Hijmans *et al.*, 2005] permettant d'obtenir une résolution de 30 secondes, soit un maillage variant de 0.8×1.3 km à 0.5×1.3 km pour la zone concernée. Les sept variables topographiques (tableau 2.5) utilisées pour la nouvelle ECNM ont été reconstruites suivant la même grille selon la méthode décrite section 2.3.2.

Affiner la résolution des variables environnementales nous permet l'interprétation de certaines données topographiques comme la pente ou l'altitude. En effet elles étaient peu exploitables, car difficile à interpréter avec un maillage de plusieurs dizaines de kilomètres. La résolution plus précise permet également d'observer les micro-variations climatiques.

Dans un effort d'objectivité, nous avons également choisi de multiplier les techniques de modélisation.

10.3.1.2 Emploi de 10 algorithmes prédictifs et consensus pondérés

Dans la première étude, les reconstructions des ECNM avaient été conduites avec MaxEnt et DesktopGARP. Les niches modélisées issues de chacun de ces algorithmes avaient de fortes similitudes. Ainsi, l'interprétation n'a pas suscité d'ambiguïté. Ne possédant aucun moyen *in vivo* de tester les modèles utilisés, il est donc difficile de juger quel est l'outil le mieux adapté à la reconstruction des niches de ces différentes cultures du Néolithique. Conscients de ce biais, nous optons donc, dans un effort d'objectivité, et afin d'optimiser l'étude, pour l'emploi de 14 techniques de modélisation (liste tableau 10.1). Douze algorithmes ont été utilisés avec leurs réglages par défaut (*Open Modeller*) plus trois algorithmes avec des réglages leur donnant moins de restrictions dans leur prédictions (nous les qualifions d'optimistes). Par la suite, les résultats sont présentés sous forme de consensus élitistes pondérés (voir chapitre 6). Les prédictions de niches sélectionnées pour réaliser ces derniers doivent avoir une AUC supérieure à 0.95 puis la pondération des prédictions retenues s'effectue selon la valeur de leur *partial ROC ratio*.

Tableau 10.1 : Sélection et pondération des algorithmes utilisés pour la reconstruction des niches écoculturelles du Néolithique ancien. Seuls les algorithmes avec une AUC supérieure à 0.95 sont retenus ; la pondération a été effectuée selon le *partial ROC ratio* (w). Les valeurs de *partial ROC ratio* inférieures à 1.5 ont été ramenées à 1 afin d'accorder une importance plus faible aux prédictions concernées

	Impressa		Cardial		LBK	
	AUC	w	AUC	w	AUC	w
ANN	0.95	1.91	0.95	1.91	0.94	1.88
ANN (optimiste)	0.93	1.86	0.93	1.85	0.75	1.49
Bioclim	0.99	1.00	0.99	1.00	0.96	1.00
Chebyshev Env. dist.	0.50	1.00	0.52	1.00	0.51	1.00
DesktopGARP BS	0.99	1.00	0.99	1.00	0.96	1.00
DesktopGARP BS (optimiste)	0.93	1.00	0.98	1.00	0.94	1.00
Envelope Score	0.99	1.99	0.99	1.99	0.96	1.93
MaxEnt	0.99	1.98	0.99	1.98	0.97	1.93
GARP-OM	0.81	1.00	0.88	1.00	0.5	1.00
GARP-OM (optimiste)	0.72	1.00	0.77	1.00	0.80	1.00
GARP-OM BS	0.92	1.00	0.96	1.00	0.93	1.00
GARP-OM BS (optimiste)	0.94	1.00	0.97	1.00	0.94	1.00
Random Forests	0.90	1.80	0.92	1.83	0.95	1.90
SVM	0.50	1.00	0.50	1.00	0.50	1.00

10.3.1.3 Analyse statistique

En complément de la première étude [Antunes, 2010; Banks *et al.*, 2013a], dans laquelle nous avons mené des analyses en composantes principales (ACP) indépendantes, nous avons cette fois réalisé une ACP commune (section 7.4.1) afin d'observer les positions écologiques relatives à chacune des cultures dans l'espace environnemental. Pour réaliser l'ACP, nous avons collecté les valeurs des variables environnementales correspondant aux prédictions de présence avec une probabilité supérieure à 0.9. Nous avons choisi de placer ce seuil relativement haut afin de pouvoir apprécier au mieux les ressemblances entre les niches du Cardial et de l'Impressa. Ce seuil de 0.9 nous permet également de réduire le bruit de fond généré par le consensus (voir section 6.2.2.3), bruit qui pourrait rendre impossible l'identification par l'ACP des espaces écologiques propres aux niches des différentes cultures étudiées.

En complément de l'ACP, nous observons la distribution des prédictions de présence qui possèdent les probabilités les plus élevées à l'aide d'histogrammes pour chaque variable environnementale. Ces derniers représentent les fréquences des lieux où les cultures sont prédites comme pouvant être présentes avec une probabilité au-delà de 0.9.

10.3.2 Résultats de la nouvelle étude

Les résultats consensuels obtenus avec les données bioclimatiques sont principalement en accord avec ceux que nous avons obtenus dans [Antunes, 2010; Banks *et al.*, 2013a].

10.3.2.1 Confirmation des principaux résultats antérieurs

La comparaison des expressions géographiques des niches écoculturelles entre les deux études ne révèle pas de différences majeures. En effet, la niche de l'Impressa (figure 10.4(a)) et celle du Cardial (figure 10.4(b)) sont très similaires. Elles couvrent la côte de la Mer Noire et la majorité de l'Anatolie, des Balkans, de l'Italie, de la Péninsule Ibérique, de la Façade Atlantique et les régions côtières d'Afrique du Nord. Comme dans l'étude précédente, on peut noter une expression géographique légèrement plus étendue pour l'Impressa. Ces similitudes signalent une écologie proche et nous permettent de les regrouper en une niche reflétant la culture méditerranéenne (figure 10.4).

La Niche prédite pour le LBK (figure 10.4(d)) quant à elle, couvre une grande partie des régions baignées par le Danube et par le Pô. Elle recouvre les plaines du Nord de l'Europe ainsi qu'une grande partie des Îles Britanniques.

Enfin, nous remarquons également dans cette étude que la niche écoculturelle des cultures méditerranéennes réunies est bien distincte de celle de la culture danubienne (figure 10.8).

Les nouvelles ECNM indiquent des présences potentielles dans des régions éloignées des occurrences archéologiques utilisées pour la modélisation. Ces prédictions qui identifient les mêmes régions que dans l'étude précédente, par exemple la Grande Bretagne pour les Danubiens et l'Afrique du Nord pour les Méditerranéens, permettent de soulever des questions en rapport avec les barrières de diffusion culturelle ou technologique. En effet, les Îles Britanniques, bien que dans la niche écoculturelle des LBK, ne sont joignables que par voie maritime, or ces groupes humains n'étaient pas tournés vers la mer. À l'inverse, les Méditerranéens, pour lesquels la mer a constitué une importante voie de diffusion et d'échanges se sont peu développés sur les côtes d'Afrique du Nord pourtant prédites comme possédant des conditions environnementales favorables à leur système adaptatif.

Les nouveaux matériels et méthodes employés permettent cependant de relever des nuances dans les expressions géographiques des niches par rapport à l'ancienne étude [Antunes, 2010; Banks *et al.*, 2013a].

10.3.2.2 Subtilités relevées par la nouvelle étude dans les expressions géographiques

À première vue, les consensus des niches écoculturelles obtenus dans la nouvelle étude sont plus proches des prédictions de MaxEnt de la première étude que ceux de GARP. Nous notons une différence un peu plus marquée entre l'expression géographique de la niche du Cardial et celle de l'Impressa. En effet, si l'on exclut les prédictions inférieures à 0.5 qui peuvent être en partie dues au bruit de fond du consensus (voir section 6.2.2.3), l'étendue des fortes prédictions est bien plus importante avec notamment des prédictions de présence en Bretagne, Normandie, Grande Bretagne et Norvège. Ces régions très distantes de l'aire d'influence de l'Impressa témoignent de la largeur écologique de cette niche. Le Cardial lui, se montre beaucoup plus spécifique avec tout de même une forte probabilité de

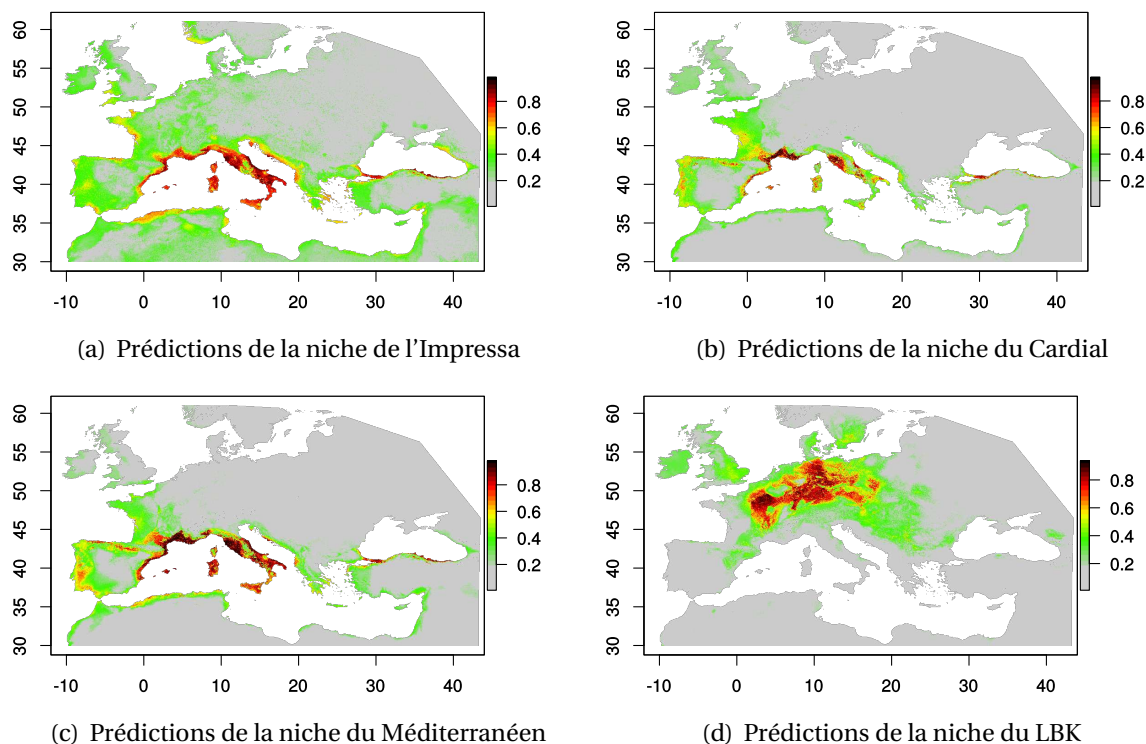


FIGURE 10.4 : Prédiction des niches écoculturelles des cultures du Néolithique

présence sur la façade atlantique, le Languedoc et le contour de la Péninsule Ibérique qui est cohérente avec l'archéologie et les théories de diffusion de ce courant culturel [Marchand et Manen, 2006].

Le consensus des Danubiens ressemble très fortement à la prédiction de MaxEnt dans [Banks *et al.*, 2013a] mais il révèle des nuances au cœur de la zone d'attribution de la culture LBK, avec notamment des prédictions de présence avec de très faibles valeurs dans le Nord-Est de la France, en Belgique et au centre de l'Allemagne.

Au-delà de la précision de l'échelle, cette étude supplémentaire a aussi été menée dans l'objectif d'apprécier les positions des niches écologiques des trois cultures les unes par rapport aux autres.

10.3.2.3 Positionnement des nouvelles niches dans l'espace écologique

L'observation des positions écologiques des prédictions de présence sélectionnées pour construire l'ACP (figure 10.5) nous permet de décrire et de comparer les niches écoculturelles en fonction des paramètres environnementaux. Nous voyons premièrement que la majorité des prédictions de présence de la niche du Cardial sont comprises dans l'espace environnemental qui convient potentiellement à la culture Impressa. Nous observons se-

Tableau 10.2 : Nombre de prédictions sélectionnées pour l'ACP. Pour conduire l'analyse, seule les prédictions de présence avec une probabilité de présence supérieure à 0.9 ont été retenues.

	Impressa	Cardial	LBK
Prédictions > 0.9	22478	24677	2892

condairement que l'espace écologique occupé par les Méditerranéens est très éloigné de celui occupé par les Danubiens.

L'ellipse correspondant à la niche de l'Impressa est celle qui présente la plus grande superficie, ceci suggère que la niche de cette culture possède la plus grande diversité de conditions environnementales. Constituant près de 45 % des prédictions sélectionnées (tableau 10.2). Elle se retrouve en position centrale dans le plan constitué des deux premières composantes de l'ACP. Néanmoins, nous relevons que cette niche semble être plus sensible que les autres aux températures minimales et à la topographie. L'inertie moyenne des prédictions de présence du Cardial se trouve elle aussi en position centrale (environ 49 % des individus de l'ACP). La taille de l'ellipse du Cardial qui synthétise un nombre d'individus quasi équivalent à celui de l'ellipse de l'Impressa (tableau 10.2), indique une forte concentration des individus. Ceci reflète des variations moins importantes des facteurs environnementaux et par conséquent une niche moins large que celle de l'Impressa.

Les niches des deux cultures méditerranéennes se superposent avec un très fort taux de recouvrement. On observe que la niche de l'Impressa est près de deux fois plus large que celle du Cardial. Ce résultat est en accord avec celui énoncé dans [Banks *et al.*, 2013a] pour les comparaisons de niches issues de la modélisation avec l'algorithme MaxEnt, mais en désaccord avec les résultats de la modélisation effectuée avec GARP. Par rapport à la niche de l'Impressa, la niche du Cardial semble se décaler suivant le premier axe, soit vers des climats plus humides en été et en hiver. Notons que ces niches ne sont pas identiques comme le laissaient supposer les résultats de l'article précédent. Les positions des ellipses laissent penser à une éventuelle sélection, voire à une réduction de niche suivant le deuxième axe de l'ACP dont la topographie et l'écart des températures journalières sont des composantes majeures. Il est intéressant de noter que la réduction et le décalage de niche modifient la position écologique de la niche du Cardial selon les deux axes de l'ACP. Ainsi, si la culture Cardiale succède à l'Impressa, la niche a alors subi une restriction et un léger décalage, entraînant un mode de vie avec un système adaptatif légèrement différent dont le reflet pourrait s'observer dans la culture matérielle. Ce déplacement vers la droite de la position écologique de cette niche la rapproche quelque peu des conditions environnementales favorables à la culture des Danubiens.

L'ellipse reflétant la position écologique de la culture LBK se trouve très éloignée de celles qui correspondent aux cultures méditerranéennes. Cette niche est principalement influencée par la quantité de précipitations des mois les plus chauds et des mois les plus froids. La superficie de l'ellipse est due principalement à un nombre d'individus sélectionnés

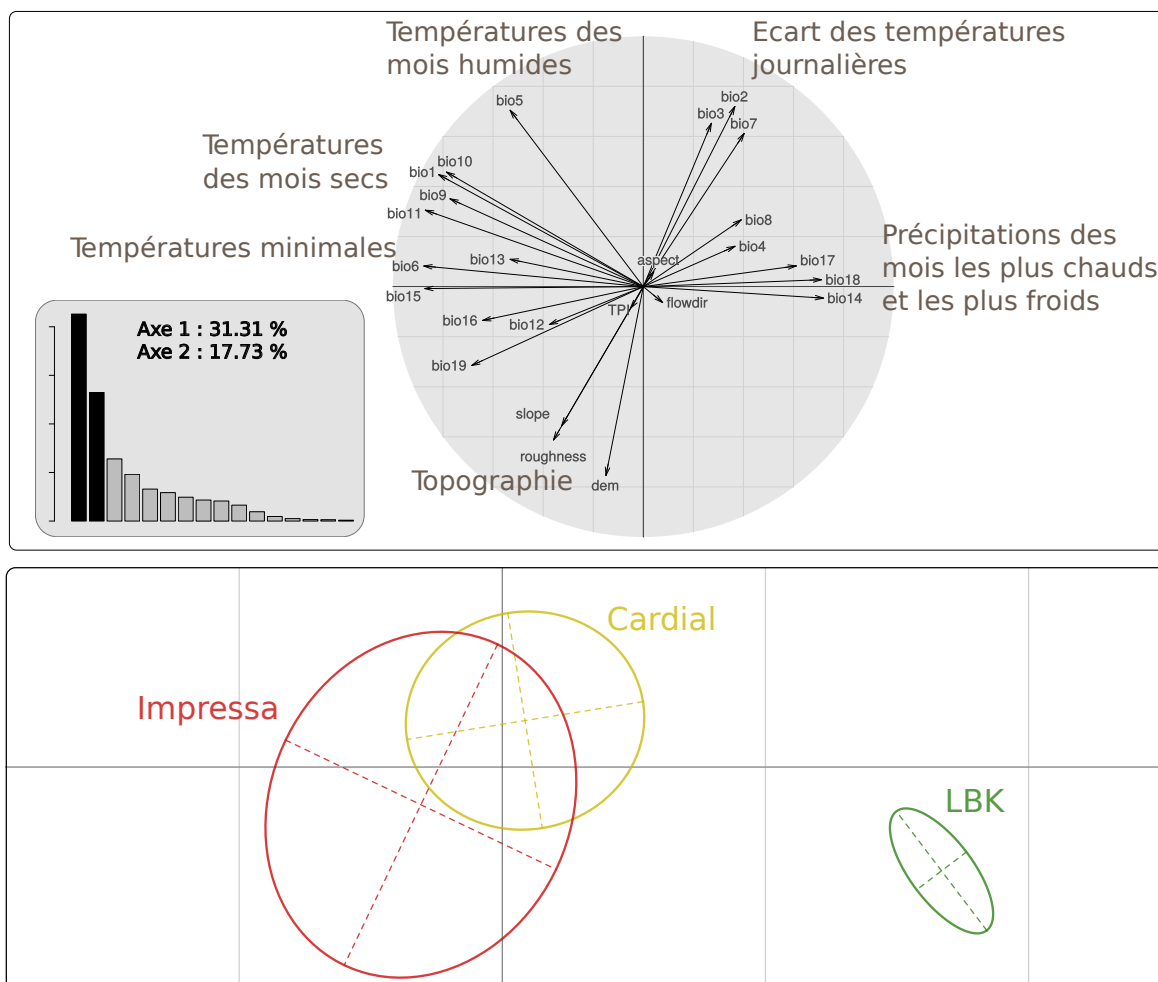


FIGURE 10.5 : Analyse en composante principale (ACP) des ECNM des cultures du Néolithique. Cette ACP a été effectuée à partir des variables environnementales prélevées en chacun des points où une présence est prédite avec une probabilité supérieure à 0.9 pour les cultures : Impressa, Cardial et LBK. Le panel du haut comprend l'histogramme des valeurs propres (chacun des bâtons représente une composante) et le cercle de corrélation des variables (voir tableaux 2.6 et 2.5). L'histogramme indique que l'analyse du plan de l'ACP constitué des deux premiers axes est suffisante pour expliquer la majorité de la variabilité du nuage de points (non représentés pour plus de lisibilité). La contribution des variables représentées par le premier axe est de 31.31% et celle par le second est de 17.73%. Le panel du bas représente les positions des niches écologiques dans le plan formé des deux premières composantes (axe 1 en abscisses, axe 2 en ordonnées). Les ellipses représentent les inerties de chacun des groupes : Impressa (en rouge), Cardial (en jaune) et LBK (en vert).

moins important pour effectuer l'analyse. L'on note cependant une forte concentration des individus dans cette région du plan, suggérant une plus grande spécialisation de cette culture par rapport aux deux autres.

Nous allons maintenant observer plus en détail les facteurs environnementaux qui composent ces niches écoculturelles.

10.3.2.4 Composition des ECN selon les variables environnementales

L'analyse des histogrammes (figures 10.6 et 10.7) nous renseigne précisément sur les conditions environnementales optimales pour chacune des cultures.

Composition des niches de l'Impressa et du Cardial. Les histogrammes de la figure 10.6 montrent la répartition des prédictions de présence des ECN de l'Impressa et du Cardial en fonction des 26 variables environnementales. Les résultats indiquent que les conditions environnementales optimales diffèrent peu, en règle générale, entre les deux cultures. On note cependant que la niche de l'Impressa possède des gammes de variation des paramètres environnementaux plus larges que celles de la niche du Cardial. La spécialisation de l'ECN du Cardial est également marquée par des fréquences plus élevées aux optima. Nous relevons des décalages dans les optima pour les facteurs environnementaux bio2, bio3 et bio7 en rapport avec les écarts thermiques journaliers, puis pour les facteurs bio6 et bio11 en relation avec les températures des périodes les plus froides. Ces décalages suggèrent une spécialisation du Cardial à un milieu plus froid, à des variations de températures journalières et annuelles plus importantes et à une saisonnalité plus marquée avec des hivers un peu plus rigoureux (température minimale du mois le plus froid : optimum pour Impressa autour de 3°C et un optimum pour Cardial à environ -1°C). La niche du Cardial se caractérise aussi par des précipitations de saison chaude plus prononcées (bio17 et bio18). On note également ce phénomène de spécialisation en observant les différences topographiques, mais ceci peut être induit par une variabilité moins importante du paysage due à l'étalement géographique plus restreint de la niche du Cardial. On note une différence très nette dans l'orientation de la pente. En effet, les Impressa sont majoritairement orientés Nord alors que les lieux correspondant à la niche du Cardial sont orientés Sud. Grâce à la haute résolution spatiale de cette étude, il est possible de discuter de l'orientation, contrairement à l'étude précédente. Cependant l'ACP dévoile que ce facteur jouait un rôle mineur dans la constitution de la niche.

Ces constats sont relativement identiques à ceux effectués dans l'étude précédente, mais la résolution plus importante ici et l'ajout de variables nous permettent de constater un décalage et une réduction de la niche du Cardial par rapport à celle de l'Impressa.

Composition des niches méditerranéennes et danubiennes. Les histogrammes de la figure 10.7 sont construits sur le même principe que ceux présentés dans la section précédente (section 10.3.2.4). Le principe est de comparer cette fois les optima des niches

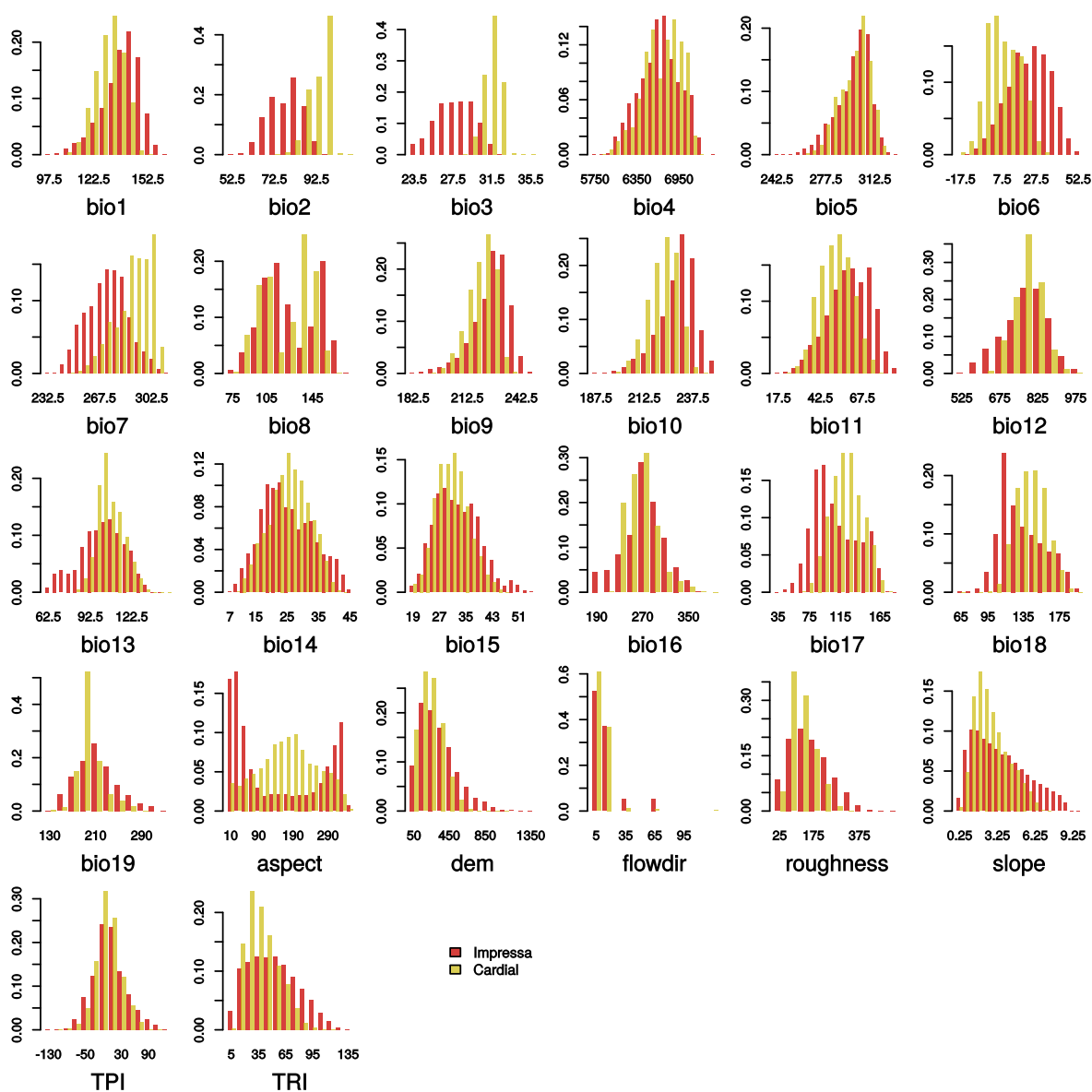


FIGURE 10.6 : Histogrammes des variables environnementales des prédictions de niches de l'Impressa et du Cardial. Chaque histogramme représente la fréquence de prédictions en fonction des valeurs de la variable environnementale. Ce graphique a été obtenu à partir des prédictions ayant une probabilité de présence supérieure à 0.9.

méditerranéennes et danubiennes. Comme vu dans l'ACP (figure 10.5) et souligné dans l'étude précédente, les niches des deux grands courants culturels du Néolithique ancien sont très distinctes. La comparaison des histogrammes figure 10.7, permet de déceler chez les Danubiens, une amplitude plus réduite des paramètres mesurés et des fréquences plus élevées au niveau des valeurs optimales ; ceci indique que ces derniers ont une niche de largeur écologique moins importante, ce qui suggère une spécialisation de niche plus importante des Danubiens que chez les Méditerranéens. Ce résultat, par ailleurs révélé par l'ACP, confirme les mesures de *breadth* et d'*overlap* que nous avons effectuées dans la première étude. Cette publication mettait en évidence la séparation des niches correspondant aux deux courants de néolithisation par l'isotherme annuel moyen de 8.5°C . Ici nous relevons un résultat identique dans l'histogramme bio1 (température moyenne annuelle) et notons le même cas de figure pour les températures des mois les plus chauds et des mois les plus froids (bio10 et bio11). Cet écart est d'autant plus marqué lorsqu'on observe les températures moyennes des mois les plus secs (bio9). Au sein du courant méditerranéen, les niches de l'Impressa et du Cardial divergent en fonction des variables en rapport avec l'amplitude thermique comme bio2 ou bio3 (voir section 10.3.2.4). La distribution des prédictions de présences du courant méditerranéen apparaît alors comme bimodale lorsque les deux niches sont réunies. Ainsi, pour ces variables, chaque optimum correspond à une culture méditerranéenne et il est intéressant de souligner que l'optimum correspondant aux conditions favorables au Cardial se superpose avec celui des LBK. Bien que les précipitations annuelles (bio12) et les précipitations des mois les plus humides (bio13) soient plus élevées dans la niche méditerranéenne, les précipitations sont plus abondantes pour les mois les plus secs (bio14) dans les régions prédites pour la culture LBK. Ceci souligne des étés moins arides dans les régions prédites comme favorables à cette culture. Enfin, on note un établissement préférentiel pour la niche des Danubiens dans des zones peu élevées et orientées au Sud. La niche des cultures méditerranéennes englobe quant à elle une variation d'altitude plus importante.

L'étude précédente montrait que les deux niches correspondant aux grandes voies de néolithisation de l'Europe ne se recouvraient qu'en de rares lieux. Grâce à la définition plus précise de notre nouvelle étude, nous pouvons observer de plus près ces zones de recouvrement où les conditions environnementales sont favorables aux deux grands courants culturels.

10.3.2.5 Chevauchement des niches méditerranéenne et danubienne

Des études sur le mobilier archéologique mettent en évidence des influences stylistiques reposant sur des voies de circulation et d'échanges établis entre les sociétés néolithiques ainsi que de possibles zones de contacts [Marchand et Manen, 2006; Otte et Noiret, 2001; Rigaud *et al.*, 2015; Sidera, 2008]

Bien que nous ayons décrit les ENM des Danubiens et des Méditerranéens comme distincts, il existe quelques zones où les niches sont prédites pour chacune des deux cultures.

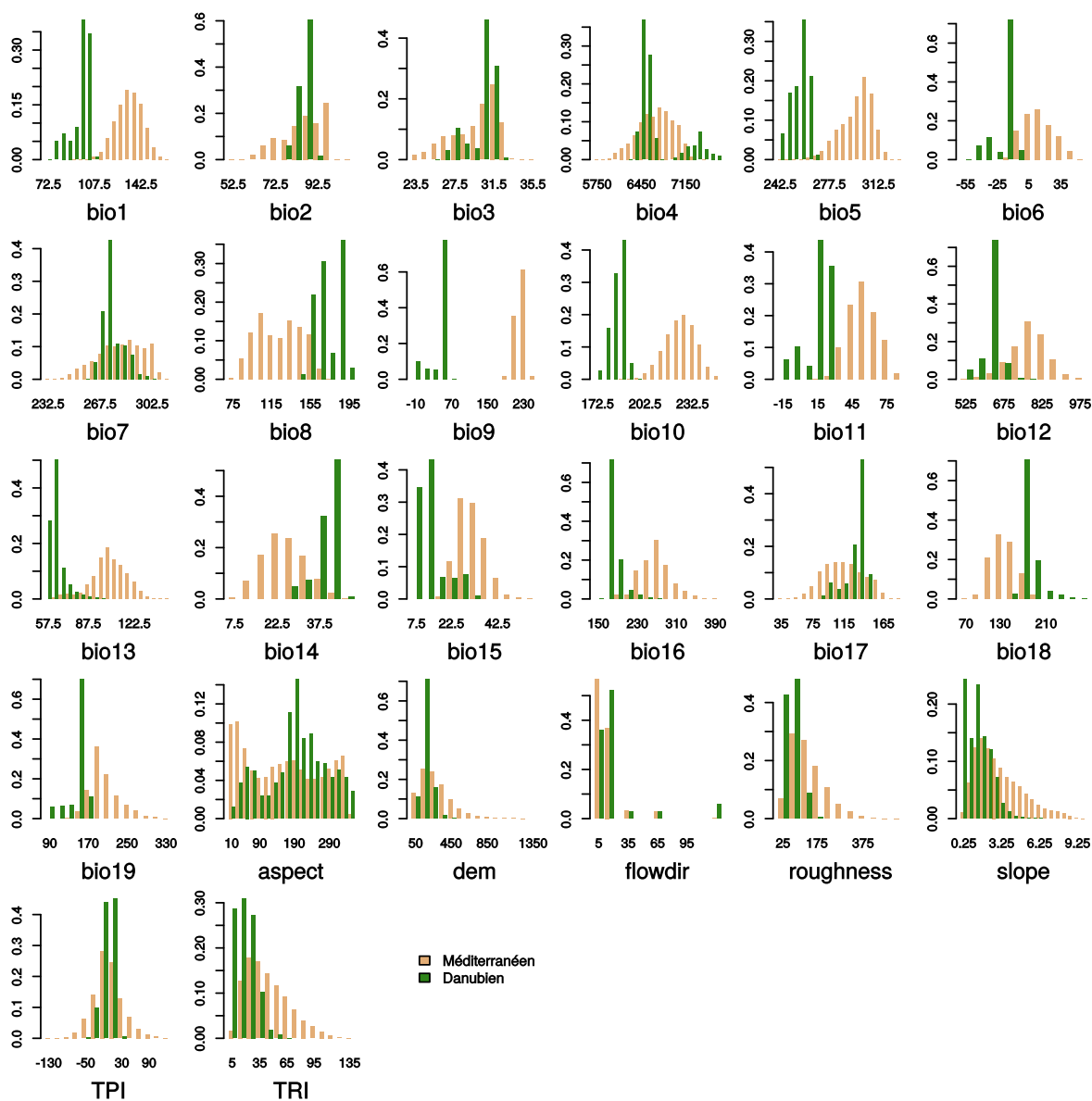


FIGURE 10.7 : Histogrammes des variables environnementales des prédictions de niches méditerranéennes et des danubiennes. Chaque histogramme représente la fréquence de prédictions en fonction des valeurs de la variable environnementale. Ce graphique a été obtenu à partir des prédictions ayant une probabilité de présence supérieure à 0.9.

Même minime, ce chevauchement de niche avait déjà été perçu par la première étude [Antunes, 2010; Banks *et al.*, 2013a]. Bénéficiant d'une résolution beaucoup plus importante avec la seconde étude, nous avons désormais la possibilité d'examiner plus en détail ces prédictions communes aux deux niches écoculturelles (figure 10.8). L'*overlap* des prédictions n'est observé que dans un bandeau de terre situé à l'Ouest de l'Europe (figure 10.8). Ce bandeau est compris entre les latitudes 38.5° et 52°N et les longitudes 1.5°O et 6°E. Dans cette portion de l'Europe, qui recouvre une partie de l'Angleterre, de la France et de l'Espagne, seules quelques zones offrent les conditions environnementales favorables aux deux cultures. Les recouvrements de niches écoculturelles les plus vastes sont localisés en France près de Vierzon, Auxerre, Brioude, Mende, de part et d'autre des Pyrénées-Orientales, et près de Guissona en Espagne.

Après une observation détaillée de ces zones, il apparaît que dans le Nord du département de l'Yonne, des sites que nous avons inclus dans l'analyse comme étant danubiens, sont localisés à proximité de lieux prédits pour chacun des courants culturels (voir zoom sur la région d'Auxerre figure 10.8). Une attribution culturelle plus précise de ces sites les affine dans des cultures de la fin du Néolithique ancien (voir données réactualisées dans [Rigaud *et al.*, 2015]). On peut relever dans cette région la présence de sites archéologiques du Néolithique moyen semblant avoir hérité de traits culturels relevant du Méditerranéen comme du Danubien. Des sites peuvent être sous l'influence d'une seule de ces deux cultures, comme Monéteau [Augereau et Chambon, 1999] (influence méditerranéenne), alors que d'autres, comme la nécropole de Gurgy [Rottier *et al.*, 2005], présentent des éléments culturels caractéristiques des deux cultures. Gurgy est le site le plus ancien connu de cette région. Le site offre une chronologie d'un millénaire débutant il y a environ 7000 ans, il a donc pu subir l'influence des deux courants de néolithisation. Confortant cette hypothèse, une étude génétique récente [Rivollat *et al.*, 2015] a relevé des fréquences d'haplotypes pouvant correspondre à la fois aux populations méditerranéennes et danubiennes.

L'exemple de la région d'Auxerre nous montre à quel point l'ECNM est capable d'identifier les lieux où règnent les conditions environnementales favorables de certaines cultures archéologiques. La modélisation de niches culturelles apporte ici un argument supplémentaire permettant de discuter de la confluence des deux courants de néolithisation dans cette région.

10.4 Perspectives apportées par la nouvelle étude

En combinant les ACP, la nouvelle étude met l'accent sur les phénomènes de conservations et de décalages de niches. Nous avons pu déterminer ainsi que le Cardial a subi de légères modifications de son système adaptatif. La résolution spatiale à 30 secondes offre la possibilité de mener des études locales. Modéliser les niches des cultures de la fin du Néolithique ancien et du début du Néolithique moyen de l'Ouest de l'Europe permettrait d'apprécier leurs positions écologiques respectives et de les comparer à celles des niches

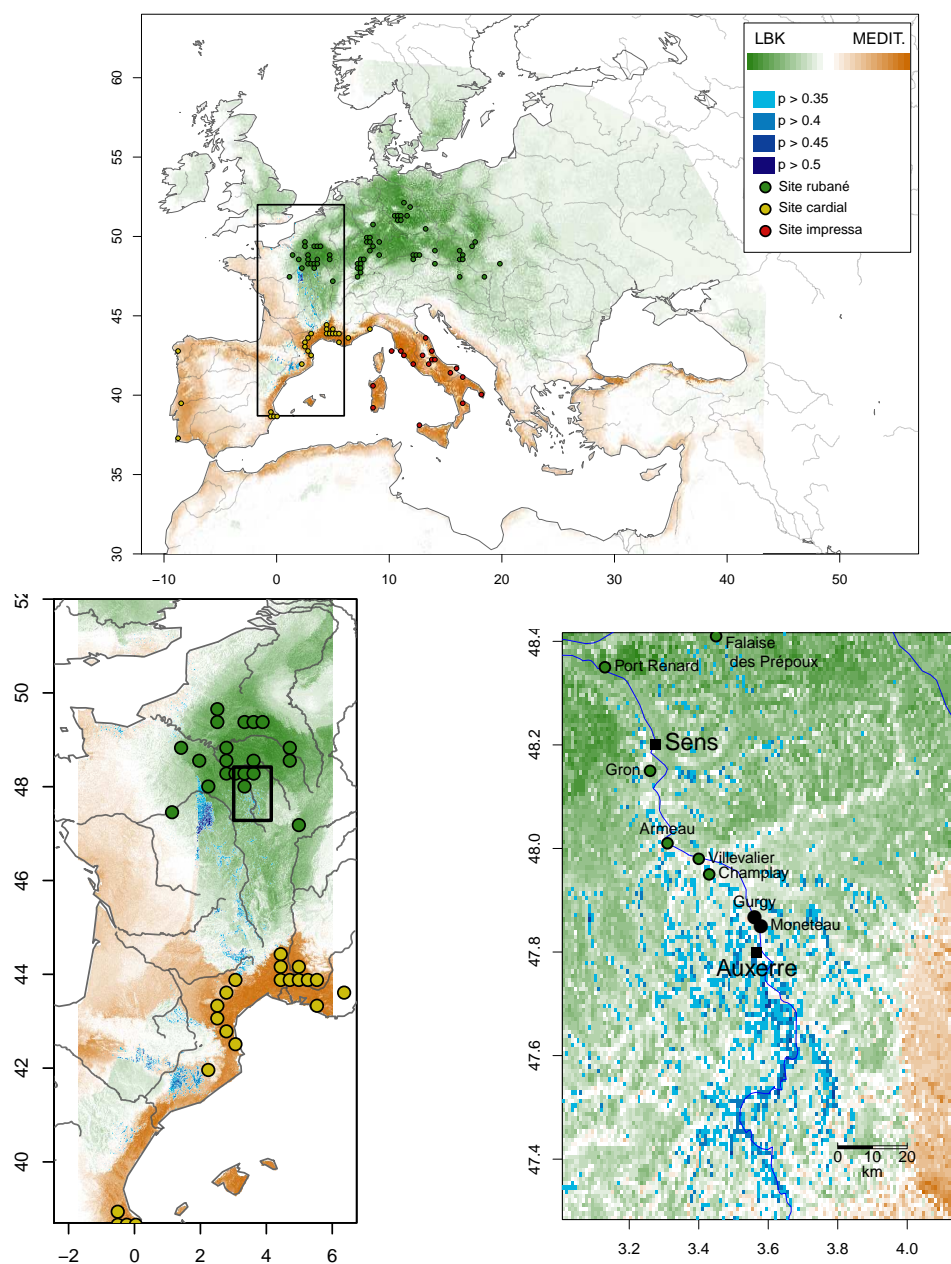


FIGURE 10.8 : Superposition des niches danubiennes et méditerranéennes. En vert est représentée la prédiction de la niche écoculturelle danubienne (LBK) et en orange la niche méditerranéenne. La superposition des deux niches écoculturelles est représentée en bleu, l'intensité du bleu varie en fonction de la probabilité de présence prédite pour chacune des niches, ainsi le bleu le plus clair correspond à l'*overlap* des prédictions avec une probabilité de présence supérieure à 0.35 pour chacune des deux niches. Le premier panel (haut) couvre toute la zone d'étude alors que le second (bas gauche) est recentré sur les prédictions communes aux deux niches (entre les méridiens -1° et 4.5° et les parallèles 38.5° et 52°). Le troisième panel est un zoom sur la région d'Auxerre où des études archéologiques ont mis en évidence la présence des deux cultures.

des deux courants précédents. Une telle étude nous apporterait des informations quant aux interactions entre les différentes sociétés de l'époque.

Le climat de l'Holocène n'est pas régulier tout au long de la période, les simulations que nous avons utilisées sont ponctuelles dans le temps, elles représentent le climat d'il y a 6000 ans. Cette date correspond à un optimum climatique [Rossignol-Strick, 1999]. Reconstruire les niches écologiques avec des pas de temps de l'ordre du siècle à partir de cette date permettrait de reconstruire la dynamique des niches des différentes cultures qui composent le Néolithique moyen de cette région et définir si le morcellement culturel qui suit la néolithisation est principalement dû aux conditions environnementales.

Synthèse du chapitre : Modélisation des niches écoculturelles des cultures initiales du Néolithique ancien

Ce chapitre présente une application de la méthodologie mise au point dans cette thèse aux cultures initiales du Néolithique.

Nous avons défini grâce à l'ECNM que les grandes voies de néolithisation méditerranéenne et danubienne s'étaient développées dans deux environnements distincts. La reconstruction des ECNM a permis l'identification des challenges environnementaux auxquels ces sociétés en transition économique ont dû faire face.

La modélisation ouvre la discussion sur la part du culturel et de l'environnemental, dans la diffusion des cultures ainsi que des adaptations et des conservations des traits culturels lors de l'établissement des populations. Ainsi, nous avons observé une spécialisation précoce des Danubiens à leur environnement et plus tardive des Méditerranéens avec le Cardial. Cette dernière est mise en évidence par une forte réduction et un léger décalage de sa niche écoculturelle par rapport à l'Impressa. Ce décalage laisse penser à un phénomène de pseudo-spéciation entre les deux cultures méditerranéennes. La pseudo-spéciation est également suggérée par le morcellement culturel se manifestant au début du Néolithique moyen. La réalisation d'ECNM pour ces différentes cultures pourraient nous indiquer si l'augmentation de la diversité culturelle est expliquée par des variations climatiques ou par des influences culturelles (comme l'apparition de spécialisations locales, l'accentuation du sédentarisme, ou l'acquisition d'autonomie).

Finalement, nous avons vu que l'utilisation de données environnementales à très haute résolution apporte des éléments de discussion quant aux zones d'éventuels contacts entre courants culturels.



Conclusions et perspectives

Synthèse générale

Nous nous sommes intéressés dans cette recherche aux relations Hommes/environnements. Les travaux de [Binford, 2001; Cavalli-Sforza *et al.*, 1988; Darwin, 1859], pour ne citer qu'eux, énoncent de possibles impacts de l'environnement sur l'histoire des peuplements, les diverses stratégies de subsistance ainsi que sur les diversités culturelles et linguistiques. Comme nous l'avons constaté, la modélisation de niches écoculturelles peut apporter des éléments de réponses à ces thématiques ; car elle permet de vérifier dans quelle mesure des cultures identifiées au sein du registre archéologique sont le reflet des conditions environnementales.

Dans ce travail doctoral, nous nous sommes attachés à reconstruire des niches écoculturelles afin de déterminer et de quantifier la part de l'environnement dans l'établissement des groupes humains étudiés. Pour mener à bien nos recherches nous nous sommes en premier lieu interrogés sur la méthode de modélisation de niches, à savoir quel était l'espace écologique modélisé et l'impact de la qualité des données d'occurrences, sur les données environnementales ainsi que sur les techniques de modélisations elles-mêmes. Puis, en second lieu, nous avons validé cette méthodologie en l'appliquant à trois cas d'études portant sur diverses populations dont les affiliations culturelles et les environnements étaient connus avec des degrés de certitudes différents.

Optimisation de l'ECNM

Nous avons vu que l'utilisation de consensus élitistes pondérés peut pallier dans une certaine mesure aux biais induits par la taille de l'échantillon, l'absence de données biotiques, la fiabilité des occurrences ainsi que le caractère stochastique des prédictions. Face

au manque de données environnementales à très haute résolution pour certaines régions à des époques spécifiques, nous avons développé une méthode de *downscaling* permettant d'obtenir de telles données à partir de simulations climatiques globales couvrant de longues périodes. Cette méthode permet d'une part l'obtention de données spécifiques à une étude et, d'autre part, de produire des couches de données environnementales qui correspondent à de multiples périodes. Ce second apport permet une analyse de la dynamique de la niche au cours du temps. Ainsi, la mise en exergue de processus évolutifs comme la réduction, l'expansion ou le décalage de niches est alors envisageable.

Démarche ethnoarchéologique

Les divers travaux menés dans cette thèse ont suivi une approche ethnoarchéologique dans le sens où nous appuyons notre raisonnement sur des données actuelles, précises et vérifiables, pour valider notre méthode et l'appliquer ensuite à des périodes anciennes.

Le premier cas d'étude porte sur la relation diversité linguistique/environnement en Nouvelle-Guinée (chapitre 8). Il constitue un référentiel actualiste puisque les conditions environnementales et les attributions culturelles sont parfaitement connues et documentées. La modélisation de niches écolinguistiques de Nouvelle-Guinée nous a permis de vérifier que la complexité de la géographie culturelle d'une région est liée à l'histoire des populations ainsi qu'aux adaptations à l'environnement. Cette étude souligne l'impact du risque écologique sur la diversité linguistique et permet de comparer différents systèmes adaptatifs afin de discuter des interactions entre les groupes linguistiques.

Le deuxième cas d'étude porte sur la disparition des colonies scandinaves subarctiques suite à l'avènement du Petit Âge Glaciaire (chapitre 9). Ce cas d'étude constitue un référentiel historique de qualité car l'histoire et les mœurs de cette société sont relativement bien documentés [Boyer, 1992; Gad, 1970]. De plus, il montre que l'impact de changements climatiques sur des populations humaines du passé peut être évalué grâce à l'ECNM. En effet, nous avons montré que la détérioration climatique a joué un rôle principal dans la disparition de la culture viking au Groenland. L'observation fine de la variation des niches, siècle par siècle, a mis en évidence une capacité de résilience suggérant l'importance des échanges entre les colonies pour leur maintien.

Les deux premiers cas d'études constituent des témoins et permettent de discuter, d'un point de vue écologique et culturel, des relations Hommes/environnements et de leur impact sur l'économie, l'histoire du peuplement ou les interactions sociales. S'inscrivant dans la continuité, le dernier cas d'étude présente une application de l'ECNM à la Préhistoire récente.

Le troisième cas d'étude établit que les populations ayant participé à la néolithisation en Europe ont développé deux systèmes adaptatifs économiques et culturels distincts qui ont diffusé dans des enveloppes environnementales qui leurs étaient propres.

Apports de l'ECNM

Les travaux conduits dans cette thèse permettent :

- d'identifier et de quantifier les paramètres environnementaux favorables à l'établissement de chaque culture ;
- d'identifier la permissivité culturelle d'un environnement ;
- d'identifier et de quantifier les risques écologiques ;
- de comparer les affinités écologiques de groupes humains ;
- de relever des régions où une compétition pour les ressources pourrait exister ;
- d'identifier des barrières environnementales ;
- de mettre en évidence l'existence de barrières culturelles ;
- d'observer l'évolution des systèmes adaptatifs.

Tous ces éléments nous aident à discuter de l'établissement de cultures dans des lieux donnés, de leur maintien, de leur abandon et de leur l'évolution.

Discussion et perspectives

Diversité culturelle

L'ECNM confirme l'hypothèse de la part de l'environnement dans l'étendue d'un territoire. Plus le risque écologique est élevée plus une population devra étendre son territoire pour accéder aux ressources. Dans une telle situation, des interactions sociales au sein de l'aire écologique correspondante au système adaptatif confèrent des avantages. Une diversité linguistique faible indique que le degré d'intercompréhension est maintenu élevé dans de tels environnements. Les barrières dressées entre les cultures sont, dans ce cas, de nature environnementale. Inversement, lorsque l'environnement est permissif à la dérive des comportements, les frontières mises en place sont d'ordre culturel. En effet, au sein d'une même niche, les humains tendent à établir des barrières de façon arbitraire (pseudo-spéciation, [Erikson, 1996]) dont les critères peuvent évoluer très rapidement. Ainsi, par des processus écologiques ou sociaux, on assiste à des morcellements culturels.

Ces morcellements peuvent s'effectuer selon des vitesses différentes. En effet, l'environnement peut rester stable sur plusieurs générations alors que des événements historiques peuvent survenir. Il est également envisageable d'imaginer que la difficulté de franchir une barrière culturelle ou une barrière environnementale n'est pas la même. Pareillement, nous pouvons concevoir que les barrières culturelles peuvent, dans certains cas, se transformer plus rapidement et qu'elles sont dotées d'une mobilité plus accrue. Identifier la nature de ces barrières permet ainsi des réflexions sur les stabilités et successions des systèmes adaptatifs observés dans le registre archéologique. Certains bagages culturels communs a plusieurs

cultures peuvent se maintenir après la disparition d'une ou de plusieurs d'entre elles. Ainsi, identifier si des morcellements de cultures ont eu lieu, par le passé, avant un événement (climatique ou historique) peut expliquer le maintien ou la disparition de cultures.

L'étude de la mise en place du Néolithique moyen du Sud-Ouest de l'Europe pourrait être un bon cas d'étude pour explorer le phénomène de morcellement de la culture danubienne et des influences méditerranéennes. Comparé au Néolithique ancien, cette période montre une haute diversité culturelle, exprimée par une multiplicité des productions matérielles, des savoirs faire et des pratiques symboliques, notamment funéraires [Beyneix, 2003; Demoule *et al.*, 2007; Jeunesse, 1997, 2004]

À partir de simulations globales couvrant les derniers 6000 ans (par exemple [Fischer et Jungclaus, 2011]) et appliquant la méthode décrite chapitre 4, il serait possible d'obtenir des données environnementales à très haute résolution pour des périodes ciblées. Couplées aux données archéologiques, celles-ci permettraient de modéliser les niches de chaque technocomplexe du Néolithique moyen. L'analyse de ces niches écoculturelles permettrait d'explorer les notions de risque écologique et de diffusion culturelle, telles que développées dans le chapitre 8 de cette thèse. La comparaison des niches obtenues avec les projections des niches prédites pour les premières cultures du Néolithique (chapitre 10) dans ces périodes plus récentes permettrait d'apprécier le décalage de niches entre les cultures du Néolithique ancien et moyen. Cette comparaison ouvrirait la discussion sur les transmissions, conservations et évolutions de ces systèmes adaptatifs.

Évolution, compétition et peuplement

Les projections de niches écoculturelles d'une culture dans différentes phases climatiques permettent d'apprécier des variations (contraction, expansion) et d'éventuels décalages de niches. La comparaison des expressions géographiques de ces niches met en exergue les conservations ou les modifications de niches. Ainsi, nous avons mis en évidence l'impact du PAG provoquant un goulot d'étranglement puis une disparition totale de la culture Viking des régions nordiques (chapitre 9). Après l'analyse des variations séculaires des niches, nous suggérons une capacité de résilience des Vikings assurée par les échanges entre les colonies qui n'a pu se produire à nouveau après la disparition de la colonie d'Islande et de celle de l'Ouest du Groenland. Cependant, pour expliquer ce phénomène, nous pouvons également évoquer la compétition avec d'autres populations. La réalisation des niches Inuits de ces périodes permettraient d'approfondir ce point.

Le peuplement de l'île principale de Nouvelle-Guinée peut également s'expliquer par la compétition. En effet, nous avons mis en évidence que les terres les moins convoitées par les Papous étaient peuplées par des Austronésiens. Nous émettons l'hypothèse d'un peuplement par opportunisme de ces derniers en ces territoires, certainement défendus avec moins d'ardeur par les populations papoues.

Cette situation peut être transposée à d'autres périodes, quand des populations migrent et viennent s'installer en des terres déjà peuplées. Il peut être imaginé, par exemple, que

la circulation des premières communautés agricoles en Méditerranée occidentale ait principalement eu lieu par cabotage et qu'elles se soient installées dans un premier temps en suivant une stratégie opportuniste. La reconstruction des niches des derniers chasseurs collecteurs du Mésolithique aboutirait à l'identification de régions plus ou moins convoitées par ces derniers et permettrait de tester cette hypothèse. Par la suite, cette étude pourrait être approfondie par la détection de zones de ralentissement de la diffusion des cultures néolithiques afin de voir si leurs vitesses de propagations sont dépendantes de contraintes environnementales ou de degrés d'acceptation de la néolithisation.

Ethnicité des populations archéologiques

Lors de ce travail nous avons modélisé les niches écoculturelles de populations dont nous étions certains des affiliations culturelles afin de vérifier et de calibrer les modélisations de niches. L'étude menée au chapitre 9 montre que la méthode développée pour la reconstruction de niches écologiques à fines résolutions spatiale et temporelle est capable d'estimer l'impact des changements environnementaux sur des populations humaines. Ainsi, pour l'étude d'une population archéologique, cette méthode apporte des éléments de discussion sur la contribution des facteurs sociaux ou écologiques responsables d'un changement de système adaptatif.

L'étude portant sur la diversité linguistique de Nouvelle-Guinée (chapitre 8) constitue un référentiel actualiste dans lequel nous pouvons établir des relations entre ethnicité et environnement. Nous envisageons de reconstituer des ECN à partir de matériel ethnographique observé dans le cadre du projet *Ethnoparure* et de les confronter aux ELN. Cette comparaison nous permettra d'identifier des types d'objets (ou des assemblages) comme étant des marqueurs d'ethnicité que nous pourrions rechercher par la suite dans le registre archéologique de la région.



Bibliographie

- ADDERLEY, W. P. et SIMPSON, I. A. (2006). Soils and palaeo-climate based evidence for irrigation requirements in Norse Greenland. *Journal of Archaeological Science*, 33(12): 1666–1679. Cité page [194](#).
- ADDERLEY, W. P., SIMPSON, I. A. et VÉSTEINSSON, O. (2008). Local-scale adaptations : A modeled assessment of soil, landscape, microclimatic, and management factors in Norse home-field productivities. *Geoarchaeology*, 23(4):500–527. Cité pages [194](#) et [196](#).
- ALLEN, J. (1996). The pre-Austronesian settlement of island Melanesia : Implications for Lapita archaeology. *Transactions of the American philosophical society*, pp. 11–27. Cité page [140](#).
- ALLEN, J. (2000). From beach to beach : The development of maritime economies in prehistoric Melanesia. *East of Wallace's Line : Studies of Past and Present Maritime Cultures of the Indo-Pacific Region*, pp. 139–176. Cité page [144](#).
- ALLOUCHE, O., TSOAR, A. et KADMON, R. (2006). Assessing the accuracy of species distribution models : prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6):1223–1232. Cité pages [76](#), [77](#), [78](#) et [110](#).
- AMANTE, C. et EAKINS, B. W. (2009). *ETOPO1 1 arc-minute global relief model : procedures, data sources and analysis*. US Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Geophysical Data Center, Marine Geology and Geophysics Division. Cité pages [50](#), [63](#) et [230](#).
- AMARASEKARE, P. (2003). Competitive coexistence in spatially structured environments : a synthesis. *Ecology Letters*, 6(12):1109–1122. Cité page [27](#).

- AMMERMAN, A. J. (2011). The paradox of early voyaging in the Mediterranean and the slowness of the Neolithic transition between Cyprus and Italy. *The Seascapes in Aegean Prehistory. Monographs of the Danish Institute at Athens*, 14:31–49. Cité page [232](#).
- AMMERMAN, A. J. et CAVALLI-SFORZA, L. L. (1971). Measuring the rate of spread of early farming in Europe. *Man*, pp. 674–688. Cité page [226](#).
- AMMERMAN, A. J., PINHASI, R. et BÁNFFY, E. (2006). Comment on "Ancient DNA from the First European Farmers in 7500-Year-Old Neolithic Sites". *Science*, 312(5782):1875–1875. Cité page [226](#).
- ANDERSON, D. G., GILLAM, J., CAITHNESS, N., PETERSON, A. T., YERKA, S. et AMMAN, C. (2005). The Clovis Radiation and Post-Clovis Settlement in North America. In *An OMLL-NSF International Symposium on Eco-Cultural Niche Modeling : Exploring the Potential of Eco-Cultural Niche Modeling for Reconstructing the Geography of Past Human Populations*, Sept. 22–26, Les Eyzies. Cité pages [42](#) et [43](#).
- ANDERSON, D. G. et GILLAM, J. C. (2000). Paleoindian colonization of the Americas : implications from an examination of physiography, demography, and artifact distribution. *American Antiquity*, pp. 43–66. Cité page [42](#).
- ANTUNES, N. (2010). Apport des algorithmes prédictifs à la compréhension des relations Homme-Environnement. Application aux processus de néolithisation de l'Europe (non publié). Mémoire de Master, Université Paul Sabatier, Toulouse. Cité pages [228](#), [232](#), [234](#), [235](#) et [243](#).
- ANTUNES, N., BANKS, W. E. et D'ERRICO, F. (2014). Evaluating Viking eco-cultural niche variability between the Medieval Climate Optimum and the Little Ice Age : a feasibility study. In *Garcia et al. (eds), Debating Spatial Archaeology. Santander, 2012*, pp. 113–130. Cité pages [51](#), [194](#), [195](#), [196](#), [197](#) et [199](#).
- ANTUNES, N., VANHAEREN, M., SCHIEFENHÖVEL, W., BANKS, W. E., RIGAUD, S. et D'ERRICO, F. (2013). Comparing ethnolinguistic affiliation and eco-cultural niches in Papua New Guinea : Archaeological Implications. In *3rd annual Meeting of the European Society for the study of Human Evolution*, Poster Presentation Number 75, p. 28, Vienna. Cité page [51](#).
- ARAÚJO, C. B., MARCONDES-MACHADO, L. O. et COSTA, G. C. (2014). The importance of biotic interactions in species distribution models : a test of the Eltonian noise hypothesis using parrots. *Journal of Biogeography*, 41(3):513–523. Cité pages [27](#), [39](#), [68](#), [78](#) et [109](#).
- ARAÚJO, M. B., CABEZA, M., THUILLER, W., HANNAH, L. et WILLIAMS, P. H. (2004). Would climate change drive species out of reserves ? An assessment of existing reserve-selection methods. *Global Change Biology*, 10(9):1618–1626. Cité page [41](#).

- ARAÚJO, M. B. et GUIBAN, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33(10):1677–1688. Cité pages [25](#), [26](#) et [27](#).
- ARAÚJO, M. B. et LUOTO, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16(6):743–753. Cité page [28](#).
- ARAÚJO, M. B. et NEW, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22(1):42–47. Cité page [117](#).
- ARAÚJO, M. B., WHITTAKER, R. J., LADLE, R. J. et ERHARD, M. (2005). Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, 14(6):529–538. Cité pages [24](#), [26](#) et [109](#).
- ARIAS CABAL, P. (2007). Neighbours but diverse : social change in north-west Iberia during the transition from the Mesolithic to the Neolithic (5500-4000 cal BC). *Proceedings of the British Academy*, 144:53–71. Cité page [226](#).
- ARNEBORG, J. (2005). Greenland irrigation systems on a West Nordic background. In *Ruralia, Conference Ruralia, Water Management in Medieval Rural Economy*, volume 27, pp. 137–145. Cité page [194](#).
- ARNEBORG, J., HEINEMEIER, J., LYNNERUP, N., NIELSEN, H. L., RUD, N. et SVEINBJÖRNSDÓTTIR, Á. E. (1999). Change of Diet of the Greenland Vikings Determined from Stable Carbon Isotope Analysis and ^{14}C Dating of Their Bones. *Radiocarbon*, 41(2):157–168. Cité page [195](#).
- ARNEBORG, J., LYNNERUP, N. et HEINEMEIER, J. (2012a). Human diet and subsistence patterns in Norse Greenland AD c. 980-AD c. 1450 : Archaeological interpretations. *Journal of the North Atlantic*, 3:119–133. Cité pages [195](#) et [223](#).
- ARNEBORG, J., LYNNERUP, N., HEINEMEIER, J., MØHL, J., RUD, N. et SVEINBJÖRNSDÓTTIR, Á. E. (2012b). Norse Greenland dietary economy ca. AD 980-ca. AD 1450 : Introduction. *Journal of the North Atlantic*, 3:1–39. Cité pages [195](#) et [223](#).
- AUBERT, M., BRUMM, A., RAMLI, M., SUTIKNA, T., SAPTOMO, E. W., HAKIM, B., MORWOOD, M., van den BERGH, G. D., KINSLEY, L. et DOSSETO, A. (2014). Pleistocene cave art from Sulawesi, Indonesia. *Nature*, 514(7521):223–227. Cité page [141](#).
- AUGEREAU, A. et CHAMBON, P. (1999). Le site néolithique de Monéteau «sur Macherin» (Yonne). In *Actes du 24e colloque interrégional sur le Néolithique, Orléans 1999*. Cité page [243](#).

- AUSTIN, M., NICHOLLS, A. et MARGULES, C. R. (1990). Measurement of the realized qualitative niche : environmental niches of five Eucalyptus species. *Ecological Monographs*, 60(2):161–177. Cité page 26.
- BABCOCK, W. H. (1917). Markland, Otherwise Newfoundland. *Geographical Review*, 4(4):309–315. Cité pages 191 et 194.
- BAKELS, C. (2014). The first farmers of the Northwest European Plain : some remarks on their crops, crop cultivation and impact on the environment. *Journal of Archaeological Science*, 51:94–97. Cité page 231.
- BANKS, W. E. (2013). *Assessing Relationships between Human Adaptative Responses and Ecology via Eco-Cultural Niche Modeling*. Habilitation à diriger des recherches en sciences et environnement, spécialité : Préhistoire, Université de Bordeaux. Cité page 34.
- BANKS, W. E., ANTUNES, N., RIGAUD, S. et D'ERRICO, F. (2013a). Ecological constraints on the first prehistoric farmers in Europe. *Journal of Archaeological Science*, 40(6):2746–2753. Cité pages 43, 44, 51, 70, 110, 228, 230, 232, 234, 235, 236, 237 et 243.
- BANKS, W. E., AUBRY, T., D'ERRICO, F., ZILHÃO, J., LIRA-NORIEGA, A. et PETERSON, A. T. (2011). Eco-cultural niches of the Badegoulian : unraveling links between cultural adaptation and ecology during the Last Glacial Maximum in France. *Journal of Anthropological Archaeology*, 30(3):359–374. Cité pages 4, 34, 43, 51 et 70.
- BANKS, W. E., D'ERRICO, F., DIBBLE, H. L., KRISHTALKA, L., WEST, D., OLSZEWSKI, D. I., PETERSON, A. T., ANDERSON, D. G., GILLAM, J. C., MONTET-WHITE, A. *et al.* (2006). Eco-cultural niche modeling : new tools for reconstructing the geography and ecology of past human populations. *Paleoanthropology*, 4:68–83. Cité pages 4, 34, 42, 45 et 51.
- BANKS, W. E., D'ERRICO, F., PETERSON, A. T., KAGEYAMA, M. et COLOMBEAU, G. (2008a). Reconstructing ecological niches and geographic distributions of caribou (*Rangifer tarandus*) and red deer (*Cervus elaphus*) during the Last Glacial Maximum. *Quaternary Science Reviews*, 27(27–28):2568–2575. Ice Age Refugia and Quaternary Extinctions : An Issue of Quaternary Evolutionary Palaeoecology. Cité page 42.
- BANKS, W. E., D'ERRICO, F., PETERSON, A. T., KAGEYAMA, M., SIMA, A. et SÁNCHEZ-GOÑI, M.-F. (2008b). Neanderthal Extinction by Competitive Exclusion. *PLoS ONE*, 3(12):e3972. Cité pages 4, 43 et 51.
- BANKS, W. E., D'ERRICO, F., PETERSON, A. T., VANHAEREN, M., KAGEYAMA, M., SEPULCHRE, P., RAMSTEIN, G., JOST, A. et LUNT, D. (2008c). Human ecological niches and ranges during the LGM in Europe derived from an application of eco-cultural niche modeling. *Journal of Archaeological Science*, 35(2):481–491. Cité pages 4, 38, 40, 42, 43 et 51.

- BANKS, W. E., D'ERRICO, F. et ZILHÃO, J. (2013b). Human climate interaction during the Early Upper Paleolithic : testing the hypothesis of an adaptive shift between the Proto-Aurignacian and the Early Aurignacian. *Journal of Human Evolution*, 64(1):39–55. Cité pages 4, 37, 38, 43 et 90.
- BANKS, W. E., ZILHÃO, J., D'ERRICO, F., KAGEYAMA, M., SIMA, A. et RONCHITELLI, A. (2009). Investigating links between ecology and bifacial tool types in Western Europe during the Last Glacial Maximum. *Journal of Archaeological Science*, 36(12):2853–2867. Cité pages 4, 43 et 51.
- BANNING, E. B. et BYRD, B. F. (1987). Houses and the changing residential unit : domestic architecture at PPNB 'Ain Ghazal, Jordan. In *Proceedings of the Prehistoric Society*, volume 53, pp. 309–325. Cambridge Univ Press. Cité page 225.
- BARKER, G. (2006). *The Agricultural Revolution in Prehistory : Why did Foragers become Farmers ? : Why did Foragers become Farmers ?* Oxford University Press. Cité page 227.
- BARKER, S., KNORR, G., EDWARDS, R. L., PARRENIN, F., PUTNAM, A. E., SKINNER, L. C., WOLFF, E. et ZIEGLER, M. (2011). 800,000 years of abrupt climate variability. *Science*, 334(6054): 347–351. Cité page 38.
- BARLOW, L. K., SADLER, J. P., OGILVIE, A. E., BUCKLAND, P. C., AMOROSI, T., INGIMUNDARSON, J. H., SKIDMORE, P., DUGMORE, A. J. et MCGOVERN, T. H. (1997). Interdisciplinary investigations of the end of the Norse Western Settlement in Greenland. *The Holocene*, 7(4):489–499. Cité pages 195 et 196.
- BARVE, N., BARVE, V., JIMÉNEZ-VALVERDE, A., LIRA-NORIEGA, A., MAHER, S. P., PETERSON, A. T., SOBERÓN, J. et VILLALOBOS, F. (2011). The crucial role of the accessible area in Ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222(11):1810–1819. Cité pages 32 et 70.
- BEAUMONT, L. J., HUGHES, L. et POULSEN, M. (2005). Predicting species distributions : use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*, 186(2):251–270. Cité pages 41, 55 et 67.
- BELLWOOD, P. (1978). *Man's conquest of the Pacific : the prehistory of Southeast Asia and Oceania*. Collins Auckland. Cité pages 143 et 147.
- BELLWOOD, P. (1985). *Prehistory of the Indo-Malayan Archipelago*. Sydney, Academic Press. Cité page 143.
- BELLWOOD, P. (1991). The Austronesian dispersal and the origin of languages. *Scientific American*, 265(1):88–93. Cité pages 139 et 140.

- BÉRAUD-VILLARS, J. M. E. (1951). *Les Normands en Mediterranee.-Paris : Michel (1951). 361 S. 8°.* A. Michel. Cité page 189.
- BERNARDO-SILVA, J., MARTINS-FERREIRA, C., MANEYRO, R. et Ochotorena de FREITAS, T. (2012). Identification of Priority Areas for Conservation of Two Endangered Parapatric Species of Red-bellied Toads Using Ecological Niche Models and Hotspot Analysis. *Natureza & Conservação*, 10. Cité pages 67 et 91.
- BEYNEIX, A. (2003). *Traditions funéraires néolithiques en France méridionale : 6000-2200 avant J.-C.* Editions Errance. Cité page 250.
- BINFORD, L. R. (1980). Willow smoke and dogs' tails : hunter-gatherer settlement systems and archaeological site formation. *American Antiquity*, pp. 4–20. Cité page 124.
- BINFORD, L. R. (2001). *Constructing frames of reference : an analytical method for archaeological theory building using hunter-gatherer and environmental data sets.* University of California Press. Cité pages 3, 37 et 247.
- BIO, A., ALKEMADE, R. et BARENDREGT, A. (1998). Determining alternative models for vegetation response analysis : a non-parametric approach. *Journal of Vegetation Science*, 9(1):5–16. Cité page 55.
- BIRDELL, J. B. (1951). The problem of the early peopling of the Americas as viewed from Asia. *Papers on the Physical Anthropology of the American Indian*, pp. 1–68. Cité page 43.
- BLUST, R. (1999). Subgrouping, circularity and extinction : some issues in Austronesian comparative linguistics. In *Selected papers from the eighth international conference on Austronesian linguistics*, volume 1, pp. 31–94. Taipei : Academia Sinica. Cité page 147.
- BOCQUET-APPEL, J. (2008). The Neolithic demographic transition, population, pressure and cultural change. *Comparative Civilizations Review*, 58:36–49. Cité page 225.
- BOCQUET-APPEL, J.-P., DEMARS, P.-Y., NOIRET, L. et DOBROWSKY, D. (2005). Estimates of Upper Palaeolithic meta-population size in Europe from archaeological data. *Journal of Archaeological Science*, 32(11):1656–1668. Cité page 68.
- BOMHARD, B., RICHARDSON, D. M., DONALDSON, J. S., HUGHES, G. O., MIDGLEY, G. E., RAIMONDO, D. C., REBELO, A. G., ROUGET, M. et THUILLER, W. (2005). Potential impacts of future land use and climate change on the Red List status of the Proteaceae in the Cape Floristic Region, South Africa. *Global Change Biology*, 11(9):1452–1468. Cité page 41.
- BONSALL, C., LENNON, R., MCSWEENEY, K. et STEWART, C. (1997). Mesolithic and Early Neolithic in the Iron Gates : a palaeodietary perspective. *Journal of European Archaeology*, 5(1):50–92. Cité page 226.

- BOOTH, P. et MCLOUGHLIN, K. (1972). The Gerbich blood group system, especially in Melanesians. *Vox sanguinis*, 22(1):73–84. Cité page 145.
- BOWER, J. (1991). The pastoral neolithic of East Africa. *Journal of World Prehistory*, 5(1):49–82. Cité page 142.
- BOYER, R. (1992). *Les Vikings*. Plon, Paris. Cité pages 192 et 248.
- BRAAK, C. J. (1988). CANOCO—an extension of DECORANA to analyze species-environment relationships. *Plant Ecology*, 75(3):159–160. Cité page 55.
- BRAME, H.-M. R. et STIGALL, A. L. (2013). Controls on niche stability in geologic time : congruent responses to biotic and abiotic environmental changes among Cincinnatian (Late Ordovician) marine invertebrates. *Paleobiology*, 40(1):70–90. Cité page 42.
- BRAUN, A. (1996). *Medizin am Zügel der Evolution. Häufigkeit und Verteilung von Erkrankungen bei zwei Naturvölkern Neuguineas*. VWB, Verlag für Wiss. und Bildung. Cité page 145.
- BREIMAN, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32. Cité page 55.
- BREIMAN, L., FRIEDMAN, J., STONE, C. J. et OLSHEN, R. A. (1984). *Classification and regression trees*. CRC press. Cité pages 54 et 55.
- BREWER, A. M. et GASTON, K. J. (2003). The geographical range structure of the holly leaf-miner. II. Demographic rates. *Journal of Animal Ecology*, 72(1):82–93. Cité page 27.
- BROENNIMANN, O., TREIER, U. A., MÜLLER-SCHÄRER, H., THUILLER, W., PETERSON, A. T. et GUIBAN, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology Letters*, 10(8):701–709. Cité page 55.
- BROTTONS, L., THUILLER, W., ARAÚJO, M. B. et HIRZEL, A. H. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27(4):437–448. Cité page 49.
- BROWN, J. H. (1995). *Macroecology*. University of Chicago Press. Cité page 24.
- BUCKLAND, P. C., AMOROSI, T., BARLOW, L. K., DUGMORE, A. J., MAYEWSKI, P. A., MCGOVERN, T. H., OGILVIE, A. E., SADLER, J. P. et SKIDMORE, P. (1996). Bioarchaeological and climatological evidence for the fate of Norse farmers in medieval Greenland. *Antiquity*, 70(267):88–96. Cité page 189.
- BUCKLAND, P. C., EDWARDS, K. J., PANAGIOTAKOPULU, E. et SCHOFIELD, J. E. (2009). Palaeoecological and historical evidence for manuring and irrigation at Garðar (Igaliku), Norse Eastern Settlement, Greenland. *The Holocene*, 19(1):105–116. Cité page 194.

- BULOT, T. (2004). Sociolinguistique. [Ressource d'Enseignement] (Non Publié). Cité page [132](#).
- BURGER, J. et THOMAS, M. G. (2011). The palaeopopulation genetics of humans, cattle and dairying in Neolithic Europe. *Human Bioarchaeology of the Transition to Agriculture*, pp. 369–384. Cité page [228](#).
- BURGES, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167. Cité page [55](#).
- CALENGE, C. (2005). *Des outils statistiques pour l'analyse des semis de points dans l'espace écologique*. Thèse de doctorat, Lyon 1. Cité page [14](#).
- CALLAWAY, R. M., BROOKER, R., CHOLER, P., KIKVIDZE, Z., LORTIE, C. J., MICHALET, R., PAOLINI, L., PUGNAIRE, F. I., NEWINGHAM, B., ASCHEHOUG, E. T. *et al.* (2002). Positive interactions among alpine plants increase with stress. *Nature*, 417(6891):844–848. Cité page [45](#).
- CALLAWAY, R. M., NADKARNI, N. M. et MAHALL, B. E. (1991). Facilitation and interference of *Quercus douglasii* on understory productivity in central California. *Ecology*, 72(4):1484–1499. Cité page [22](#).
- CANTOR, S. B., SUN, C. C., TORTOLERO-LUNA, G., RICHARDS-KORTUM, R. et FOLLEN, M. (1999). A comparison of C/B ratios from studies using receiver operating characteristic curve analysis. *Journal of Clinical Epidemiology*, 52(9):885–892. Cité page [72](#).
- CAPELLI, C., WILSON, J. F., RICHARDS, M., STUMPF, M. P., GRATRIX, F., OPPENHEIMER, S., UNDERHILL, P., PASCALI, V. L., KO, T.-M. et GOLDSTEIN, D. B. (2001). A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. *The American Journal of Human Genetics*, 68(2):432–443. Cité pages [139](#), [143](#) et [147](#).
- CARNES, B. A. et SLADE, N. A. (1982). Some comments on niche analysis in canonical space. *Ecology*, pp. 888–893. Cité page [125](#).
- CARPENTER, G., GILLISON, A. et WINTER, J. (1993). DOMAIN : a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity & Conservation*, 2(6):667–680. Cité page [55](#).
- CASHDAN, E. (2001). Ethnic diversity and its environmental determinants : Effects of climate, pathogens, and habitat diversity. *American Anthropologist*, 103(4):968–991. Cité page [133](#).
- CAUVIN, J. (1994). *Naissance des divinités, naissance de l'agriculture : la révolution des symboles au Néolithique*. Cambridge University Press. Cité pages [226](#) et [227](#).

- CAUWE, N., DOLUKHANOV, P., KOZLOWSKI, P. et VAN BERG, P.-L. (2007). *Le néolithique en Europe*. Armand Colin. Cité page [228](#).
- CAVALLI-SFORZA, L. L., MENOZZI, P. et PIAZZA, A. (1994). *The History and Geography of Human Genes*. Princeton university press. Cité pages [35](#), [139](#), [145](#), [146](#) et [147](#).
- CAVALLI-SFORZA, L. L., PIAZZA, A., MENOZZI, P. et MOUNTAIN, J. (1988). Reconstruction of human evolution : bringing together genetic, archaeological, and linguistic data. *Proceedings of the National Academy of Sciences*, 85(16):6002–6006. Cité pages [132](#) et [247](#).
- CHAPPELL, J. (2005). Geographic changes of coastal lowlands in the Papuan past. *Papuan pasts : Cultural, linguistic and biological histories of Papuan-speaking peoples*, pp. 525–540. Cité page [141](#).
- CHASE, J. M. et LEIBOLD, M. A. (2003). *Ecological niches : linking classical and contemporary approaches*. University of Chicago Press. Cité pages [10](#) et [24](#).
- CHEN, Y. (2008). Global potential distribution of an invasive species, the yellow crazy ant (*Anoplolepis gracilipes*) under climate change. *Integrative Zoology*, 3(3):166–175. Cité page [55](#).
- CHESTERMAN, A. (1997). *Memes of translation : The spread of ideas in translation theory*, volume 22. John Benjamins Publishing. Cité page [150](#).
- CHIKHI, L., DESTRO-BISOL, G., BERTORELLE, G., PASCALI, V. et BARBUJANI, G. (1998). Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proceedings of the National Academy of Sciences*, 95(15):9053–9058. Cité page [226](#).
- CHIKHI, L., NICHOLS, R. A., BARBUJANI, G. et BEAUMONT, M. A. (2002). Y genetic data support the Neolithic demic diffusion model. *Proceedings of the National Academy of Sciences*, 99(17):11008–11013. Cité page [226](#).
- CHILDE, V. G. (1925). *The Dawn of European Civilization*. London : Knopf, Alfred A., p. 365. Cité page [226](#).
- CLARK, G. (1975). *The earlier stone age settlement of Scandinavia*. Cambridge University Press Archive. Cité page [34](#).
- CLARK, J. (1965). Radiocarbon dating and the expansion of farming culture from the Near East over Europe. *Proceedings of the Prehistoric Society (New Series)*, 31:58–73. Cité page [226](#).
- CODDING, B. F. et JONES, T. L. (2013). Environmental productivity predicts migration, demographic, and linguistic patterns in prehistoric California. *Proceedings of the National Academy of Sciences*, 110(36):14569–14573. Cité pages [134](#), [146](#) et [181](#).

- COHEN, J. (1960). Weighted kappa : Nominal scale agreement provision for scaled disagreement or partial credit. *Educational and Psychological Measurement*, 20:37–46. Cité page 77.
- COLLARD, I. E. et FOLEY, R. A. (2002). Latitudinal patterns and environmental determinants of recent human cultural diversity : do humans follow biogeographical rules ? *Evolutionary Ecology Research*, 4(3):371–383. Cité pages 37 et 133.
- COUDART, A. (1993). De l'usage de l'architecture domestique et de l'anthropologie sociale dans l'approche des sociétés néolithiques : l'exemple du Néolithique danubien. *Documents d'archéologie française*, (41):114–135. Cité page 228.
- CRAMER, J. S. (2003). *Logit models from economics and other fields*. Cambridge University Press. Cité page 72.
- CRAWFORD, G. W. (2006). East Asian plant domestication. *Archaeology of Asia*, pp. 77–95. Cité page 142.
- CRAWFORD, G. W., UNDERHILL, A., ZHAO, Z., LEE, G., FEINMAN, G., NICHOLAS, L., LUAN, F., YU, H., FANG, H. et CAI, F. (2005). Late Neolithic Plant Remains from Northern China : Preliminary Results from Liangchengzhen, Shandong1. Cité page 142.
- CRISP, M. D., ARROYO, M. T., COOK, L. G., GANDOLFO, M. A., JORDAN, G. J., MCGLONE, M. S., WESTON, P. H., WESTOBY, M., WILF, P. et LINDER, H. P. (2009). Phylogenetic biome conservatism on a global scale. *Nature*, 458(7239):754–756. Cité page 34.
- CURRIE, T. E. et MACE, R. (2012). The evolution of ethnolinguistic diversity. *Advances in Complex Systems*, 15(01n02):1150006. Cité page 174.
- CUTLER, D. R., EDWARDS JR, T. C., BEARD, K. H., CUTLER, A., HESS, K. T., GIBSON, J. et LAWLER, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11):2783–2792. Cité page 55.
- CYMBRON, T., FREEMAN, A. R., MALHEIRO, M. I., VIGNE, J.-D. et BRADLEY, D. G. (2005). Micro-satellite diversity suggests different histories for Mediterranean and Northern European cattle populations. *Proceedings of the Royal Society of London B : Biological Sciences*, 272(1574):1837–1843. Cité pages 226 et 227.
- DANSGAARD, W., JOHNSEN, S., REEH, N., GUNDESTRUP, N., CLAUSEN, H. et HAMMER, C. (1975). Climatic changes, Norsemen and modern man. *Nature*, 255:24–28. Cité pages 192, 195 et 196.
- DARWIN, C. (1859). On the origins of species by means of natural selection. *London : Murray*, p. 247. Cité page 247.

- DE MEYER, M., ROBERTSON, M., PETERSON, A. T. et MANSELL, M. (2008). Ecological niches and potential geographical distributions of Mediterranean fruit fly (*Ceratitis capitata*) and Natal fruit fly (*Ceratitis rosa*). *Journal of Biogeography*, 35(2):270–281. Cité page 55.
- DE SOUZA MUÑOZ, M. E., DE GIOVANNI, R., de SIQUEIRA, M. F., SUTTON, T., BREWER, P., PEREIRA, R. S., CANHOS, D. A. L. et CANHOS, V. P. (2011). openModeller : a generic approach to species' potential distribution modelling. *Geoinformatica*, 15(1):111–135. Cité pages 58 et 114.
- DE LONG, E. R., DE LONG, D. M. et CLARKE-PEARSON, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves : a nonparametric approach. *Biometrics*, pp. 837–845. Cité page 78.
- DELPECH, F. (1999). Biomasse d'Ongulés au Paléolithique et inférences sur la démographie/Ungulate Biomass During the Paleolithic Period and Inferences Concerning Human Demography. *Paléo*, 11(1):19–42. Cité page 37.
- DEMOULE, J.-P. et al. (2007). *La révolution néolithique en France*. Découverte. Cité page 250.
- DENNEL, R. et PORR, M. (2014). *Southern Asia, Australia and the Search for Human Origins*. Cambridge University Press. Cité page 142.
- D'ERRICO, F. et BANKS, W. E. (2013). Identifying Mechanisms behind Middle Paleolithic and Middle Stone Age Cultural Trajectories. *Current Anthropology*, 0(0):p. S000. Cité pages 4, 44 et 51.
- D'ERRICO, F., BANKS, W. E. et CLOBERT, J. (2012). Human expansion : research tools, evidence, mechanisms. *Dispersal Ecology and Evolution*, pp. 433–447. Cité pages 36 et 38.
- D'ERRICO, F. et SÁNCHEZ GOÑI, M. F. (2003). Neandertal extinction and the millennial scale climatic variability of {OIS} 3. *Quaternary Science Reviews*, 22(8–9):769–788. Cité pages 3 et 43.
- DESCOLA, P. (2004). Le sauvage et le domestique. *Communications*, 76(1):17–39. Cité page 226.
- DETTMERS, R., BUEHLER, D. et BARTLETT, J. (2002). A test and comparison of wildlife-habitat modeling techniques for predicting bird occurrence at a regional scale. *Predicting species occurrences : issues of accuracy and scale*, pp. 607–615. Cité page 55.
- DI PIAZZA, A. (1996). P. Pétrequin & A.-M. Pétrequin, Ecologie d'un outil : la hache de pierre en Irian Jaya (Indonésie). *L'Homme*, 36(137):272–273. Cité page 142.
- DIAMOND, J. M. (1988). Express train to Polynesia. *Nature*, 336:307–308. Cité page 143.

- DIAMOND, J. M. (1997). *Guns, germs and steel : a short history of everybody for the last 13,000 years*. Random House. Cité page 37.
- DIAMOND, J. M. (2005). *Collapse : How societies choose to fall or survive*. London : Penguin. Cité page 195.
- DODD, L. E. et PEPE, M. S. (2003). Partial AUC Estimation and Regression. *Biometrics*, 59(3):614–623. Cité page 78.
- DOLUKHANOV, P., SHUKUROV, A., GRONENBORN, D., SOKOLOFF, D., TIMOFEEV, V. et ZAITSEVA, G. (2005). The chronology of Neolithic dispersal in Central and Eastern Europe. *Journal of Archaeological Science*, 32(10):1441–1458. Cité page 228.
- DRAKE, J. M., RANDIN, C. et GUISAN, A. (2006). Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, 43(3):424–432. Cité page 55.
- DRAPER, N. et SMITH, H. (1981). *Applied regression analysis*. John Wiley, New York. Cité page 55.
- DRAY, S., DUFOUR, A.-B. *et al.* (2007). The ade4 package : implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22(4):1–20. Cité pages 165 et 206.
- DUDA, R. O., HART, P. E. et STORK, D. G. (2001). *Pattern classification*. 2nd. Edition. New York. Cité page 55.
- DUFRESNE, J.-L., FOIJOLS, M.-A., DENVIL, S., CAUBEL, A., MARTI, O., AUMONT, O., BALKANSKI, Y., BEKKI, S., BELLENGER, H., BENSHILA, R. *et al.* (2013). Climate change projections using the IPSL-CM5 Earth System Model : from CMIP3 to CMIP5. *Climate Dynamics*, 40(9-10):2123–2165. Cité pages 65 et 202.
- DUGMORE, A. J., CHURCH, M. J., BUCKLAND, P. C., EDWARDS, K. J., LAWSON, I., MCGOVERN, T. H., PANAGIOTAKOPULU, E., SIMPSON, I. A., SKIDMORE, P. et SVEINBJARNARDÓTTIR, G. (2005). The Norse landnám on the North Atlantic islands : an environmental impact assessment. *Polar Record*, 41(01):21–37. Cité pages 192, 193, 195 et 196.
- DUGMORE, A. J., MCGOVERN, T. H., VÉSTEINSSON, O., ARNEBORG, J., STREETER, R. et KELLER, C. (2012). Cultural adaptation, compounding vulnerabilities and conjunctures in Norse Greenland. *Proceedings of the National Academy of Sciences*, 109(10):3658–3663. Cité page 194.
- D’ANDREA, W. J., HUANG, Y., FRITZ, S. C. et ANDERSON, N. J. (2011). Abrupt Holocene climate change as an important factor for human migration in West Greenland. *Proceedings of the National Academy of Sciences*, 108(24):9765–9769. Cité page 195.

- ELITH, J. et BURGMAN, M. (2002). Predictions and their validation : rare plants in the Central Highlands, Victoria, Australia. *Predicting species occurrences : issues of accuracy and scale*, pp. 303–314. Cité page [55](#).
- ELITH, J. et LEATHWICK, J. R. (2009). Species distribution models : ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40(1):677. Cité pages [27](#) et [57](#).
- ELITH, J., LEATHWICK, J. R. et HASTIE, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4):802–813. Cité page [55](#).
- ELITH, J., PHILLIPS, S. J., HASTIE, T., DUDÍK, M., CHEE, Y. E. et YATES, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1):43–57. Cité page [57](#).
- ELTON, C. (1927). *Animal Ecology, 1927. Sidgwick and Jackson, London*. Cité page [10](#).
- ENGLER, R., GUIBAN, A. et RECHSTEINER, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41(2):263–274. Cité page [41](#).
- ERIKSON, E. H. (1985). Pseudospeciation in the nuclear age. *Political Psychology*, pp. 213–217. Cité page [132](#).
- ERIKSON, K. (1996). On pseudospeciation and social speciation. *Genocide : War and Human Survival*, pp. 51–58. Cité pages [4](#) et [249](#).
- EVANS, J. S. et CUSHMAN, S. A. (2009). Gradient modeling of conifer species using random forests. *Landscape Ecology*, 24(5):673–683. Cité page [55](#).
- FAIRBAIRN, A. S., HOPE, G. S. et SUMMERHAYES, G. R. (2006). Pleistocene occupation of New Guinea's highland and subalpine environments. *World Archaeology*, 38(3):371–386. Cité page [142](#).
- FARBER, O. et KADMON, R. (2003). Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, 160(1-2):115–130. Cité page [55](#).
- FERTIG, W., REINERS, W., SCOTT, J., HEGLUND, P., MORRISON, M. et HAUFLE, J. (2002). Predicting presence/absence of plant species for range mapping : a case study from Wyoming. *Predicting species occurrences : issues of accuracy and scale*, pp. 483–489. Cité page [55](#).
- FIELD, C. B. et VAN AALST, M. (2014). *Climate change 2014 : impacts, adaptation, and vulnerability*, volume 1. IPCC. Cité page [3](#).

- FIELDING, A. H. et BELL, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(01):38–49. Cité pages 78 et 110.
- FIELDING, A. H. et HAWORTH, P. F. (1995). Testing the generality of bird-habitat models. *Conservation Biology*, 9(6):1466–1481. Cité page 72.
- FINCHER, C. L. et THORNHILL, R. (2008). A parasite-driven wedge : infectious diseases may explain language and other biodiversity. *Oikos*, 117(9):1289–1297. Cité page 133.
- FISCHER, N. et JUNGCLAUS, J. H. (2011). Evolution of the seasonal temperature cycle in a transient Holocene simulation : orbital forcing and sea-ice. *Climate of the Past*, 7:1139–1148. Cité page 250.
- FLEMONS, P., GURALNICK, R., KRIEGER, J., RANIPETA, A. et NEUFELD, D. (2007). A web-based GIS tool for exploring the world's biodiversity : The Global Biodiversity Information Facility Mapping and Analysis Portal Application (GBIF-MAPA). *Ecological Informatics*, 2(1):49–60. Cité page 49.
- FOLEY, W. A. (1986). *The Papuan languages of New Guinea*. Cambridge University Press. Cité page 146.
- FORT, J. et MÉNDEZ, V. (1999). Time-delayed theory of the Neolithic transition in Europe. *Physical review letters*, 82(4):867. Cité page 225.
- FOX, J. (2010). Polycor : polychoric and polyserial correlations. R package version 0.7-8. Online [URL] <http://www.cran.r-project.org/web/packages/polycor/index.html> (Accessed on 31 August 2010). bibr12. Cité page 69.
- FRANKLIN, J. (1998). Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, 9(5):733–748. Cité page 55.
- FRANKLIN, J. (2009). *Mapping species distributions : spatial inference and prediction*. Cambridge University Press. Cité page 47.
- FREDSKILD, B. (1992). Erosion and vegetational changes in South Greenland caused by agriculture. *Geografisk Tidsskrift-Danish Journal of Geography*, 92(1):14–21. Cité page 194.
- FREDSKILD, B. et HUMLE, L. (1991). Plant remains from the Norse farm Sandnes in the Western Settlement, Greenland. *Acta Borealia*, 8(1):69–81. Cité page 194.

- FREI, K. M., COUTU, A. N., SMIAROWSKI, K., HARRISON, R., MADSEN, C. K., ARNEBORG, J., FREI, R., GUÐMUNDSSON, G., SINDBÆK, S. M., WOOLLETT, J. *et al.* (2015). Was it for walrus? Viking Age settlement and medieval walrus ivory trade in Iceland and Greenland. *World Archaeology*, (ahead-of-print):1–28. Cité page 195.
- FRIEDMAN, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1):1–67. Cité pages 54 et 55.
- GAD, F. (1970). *The History of Greenland. In : Earliest Times to 1700, London and Hust Company*, volume 1. Cité pages 192, 197 et 248.
- GASTON, K. J. (2003). *The structure and dynamics of geographic ranges*. Oxford University Press. Cité pages 18, 30 et 32.
- GAUSE, G. (1934). The struggle for existence. *Williams and Wilkins, Baltimore*. Cité pages 19, 21, 22 et 45.
- GAVIN, M. C. et STEPP, J. R. (2014). Rapoport's Rule Revisited : Geographical Distributions of Human Languages. *PLoS ONE*, 9(9):e107623. Cité page 133.
- GEHRING, T. M. et SWIHART, R. K. (2003). Body size, niche breadth, and ecologically scaled responses to habitat fragmentation : mammalian predators in an agricultural landscape. *Biological Conservation*, 109(2):283–295. Cité page 55.
- GENT, P. R., DANABASOGLU, G., DONNER, L. J., HOLLAND, M. M., HUNKE, E. C., JAYNE, S. R., LAWRENCE, D. M., NEALE, R. B., RASCH, P. J., VERTENSTEIN, M. *et al.* (2011). The community climate system model version 4. *Journal of Climate*, 24(19):4973–4991. Cité page 67.
- GILLAM, J. C., ANDERSON, D. G., TOWNSEND, A. et PETERSON, A. T. (2007). A Continental-scale Perspective on the Peopling of the Americas : Modeling Geographic Distributions and Ecological Niches of Pleistocene Populations. *Current Research in the Pleistocene*, 24:86–90. Cité pages 42 et 43.
- GILLAM, J. C. et TABAREV, A. (2006). Geographic information systems and predictive modeling : Prospects for Far East archaeology. *Archaeological Elucidation of the Japanese Fundamental Culture in East Asia. Kokugakuin Univeristy 21st Century COE Program, Tokyo, Archaeology Series*, 7:63–76. Cité pages 42 et 43.
- GOLDING, K. A., SIMPSON, I. A., SCHOFIELD, J. E. et EDWARDS, K. J. (2011). Norse–Inuit interaction and landscape change in southern Greenland? A geochronological, pedological, and palynological investigation. *Geoarchaeology*, 26(3):315–345. Cité page 223.
- GOLSON, J. (1976). Archaeology and agricultural history in the New Guinea Highlands. *In : Sieveking, G. de, Longworth, L.A. & Wilson, K.E. (eds.) Problems in Economic and Social Archaeology*, pp. 201–220. Cité page 142.

- GOLSON, J. (1991). [Transitions to Agriculture in the Pacific Region] The New Guinea Highlands on the eve of agriculture. *Bulletin of the Indo-Pacific Prehistory Association*, 11:82–91. Cité page [142](#).
- GOLSON, J. (2005). Introduction to the chapters on archaeology and ethnology. *Pawley, Attenborough, Golson and Hide*, pp. 221–233. Cité page [146](#).
- GORDON, R. G. et GRIMES, B. F. (2005). *Ethnologue : Languages of the world*, volume 15. Sil international Dallas, TX. Cité page [147](#).
- GÖTHERSTRÖM, A., ANDERUNG, C., HELLBORG, L., ELBURG, R., SMITH, C., BRADLEY, D. G. et ELLEGREN, H. (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society of London B : Biological Sciences*, 272(1579):2345–2351. Cité page [226](#).
- GRAVIER, M. (1955). *La saga d'Eric le Rouge : Le récit des Groenlandais*. Aubier. Cité page [189](#).
- GRAY, R. D. et ATKINSON, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*, 426(6965):435–439. Cité page [226](#).
- GRAY, R. D., DRUMMOND, A. J. et GREENHILL, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, 323(5913):479–483. Cité page [144](#).
- GREEN, R. C. (2000). Lapita and the cultural model for intrusion, integration and innovation. *Australian archaeologist : collected papers in honour of Jim Allen*, pp. 372–92. Cité pages [143](#) et [144](#).
- GRINNELL, J. (1904). The Origin and Distribution of the Chest-Nut-Backed Chickadee. *The Auk*, 21(3):364–382. Cité pages [10](#) et [19](#).
- GRINNELL, J. (1917). The Niche-Relationships of the California Thrasher. *The Auk*, 34(4):427–433. Cité pages [10](#) et [15](#).
- GRONENBORN, D. (2003). Migration, acculturation and culture change in western temperate Eurasia, 6500–5000 cal BC. *Documenta Praehistorica*, 30:79–91. Cité page [3](#).
- GUIASU, S. et SHENITZER, A. (1985). The principle of maximum entropy. *The Mathematical Intelligencer*, 7(1):42–48. Cité page [55](#).
- GUILAINE, J. (2003). *De la vague à la tombe. La conquête néolithique de la Méditerranée*. Seuil, Paris. Cité pages [226](#) et [227](#).
- GUISAN, A. et THUILLER, W. (2005). Predicting species distribution : offering more than simple habitat models. *Ecology Letters*, 8(9):993–1009. Cité pages [24](#), [26](#) et [90](#).

- GUISAN, A., WEISS, S. B. et WEISS, A. D. (1999). GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology*, 143(1):107–122. Cité page 55.
- GUISAN, A. et ZIMMERMANN, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2–3):147–186. Cité page 26.
- GUO, Q. et LIU, Y. (2010). ModEco : an integrated software package for ecological niche modeling. *Ecography*, 33(4):637–642. Cité pages 55 et 58.
- HAAK, W., BALANOVSKY, O., SANCHEZ, J. J., KOSHEL, S., ZAPOROZHCHENKO, V., ADLER, C. J., DER SARKISSIAN, C. S. I., BRANDT, G., SCHWARZ, C., NICKLISCH, N., DRESELY, V., FRITSCH, B., BALANOVSKA, E., VILLEMS, R., MELLER, H., ALT, K. W., COOPER, A. et the GENOGRAPHIC CONSORTIUM (2010). Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLoS Biology*, 8(11):e1000536. Cité page 226.
- HAAK, W., FORSTER, P., BRAMANTI, B., MATSUMURA, S., BRANDT, G., TÄNZER, M., VILLEMS, R., RENFREW, C., GRONENBORN, D., ALT, K. W. *et al.* (2005). Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science*, 310(5750):1016–1018. Cité page 226.
- HABERLE, S. G. (2003). The emergence of an agricultural landscape in the highlands of New Guinea. *Archaeology in Oceania*, 38(3):149–158. Cité page 142.
- HABERLE, S. G., HOPE, G. S. et VAN DER KAARS, S. (2001). Biomass burning in Indonesia and Papua New Guinea : natural and human induced fire events in the fossil record. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 171(3):259–268. Cité page 142.
- HALL, L. S., KRAUSMAN, P. R. et MORRISON, M. L. (1997). The habitat concept and a plea for standard terminology. *Wildlife Society Bulletin*, pp. 173–182. Cité pages 10 et 14.
- HARVEY, P. H. et PAGEL, M. D. (1991). *The comparative method in evolutionary biology*, volume 239. Oxford university press Oxford. Cité page 34.
- HASSAN, F. A. (1981). Demographic archaeology. *Studies in Archaeology*. Cité page 68.
- HASTIE, T. et TIBSHIRANI, R. (1990). *Generalized additive models*. London : Chapman and Hall. Cité pages 54, 55 et 92.
- HASTIE, T. et TIBSHIRANI, R. (1996). Discriminant analysis by Gaussian mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 155–176. Cité page 55.
- HAYS, T. E. (1994). *Encyclopedia of World Cultures : Oceania*, volume 2. G.K and Co. Hall. Boston, Massachusetts. Cité page 142.
- HEIBL, C. et CALENGE, C. (2013). *phyloclim : Integrating Phylogenetics and Climatic Niche Modeling*. R package version 0.9-4. Cité pages 125 et 176.

- HEIKKINEN, R. K., LUOTO, M., VIRKKALA, R., PEARSON, R. G. et KORBER, J.-H. (2007). Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, 16(6):754–763. Cité page [28](#).
- HENRARD, D. (2003). Le Mésolithique du bassin de l'Ourthe (Belgique) : implantation dans le paysage et néolithisation. *L'Anthropologie*, 107(5):615–644. Cité page [226](#).
- HENRICH, J. (2004). Demography and cultural evolution : how adaptive cultural processes can produce maladaptive losses : the Tasmanian case. *American Antiquity*, pp. 197–214. Cité page [40](#).
- HERNANDEZ, P. A., GRAHAM, C. H., MASTER, L. L. et ALBERT, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29(5):773–785. Cité pages [54](#), [60](#), [61](#) et [110](#).
- HEYLIGHEN, F. (1998). What makes a meme successful? Selection criteria for cultural evolution. *Association Internationale de Cybernetique*. Cité page [150](#).
- HIJMANS, R. J., CAMERON, S. E., PARRA, J. L., JONES, P. G. et JARVIS, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978. Cité pages [50](#), [67](#), [68](#), [91](#), [150](#) et [233](#).
- HIJMANS, R. J., PHILLIPS, S., LEATHWICK, J. et ELITH, J. (2013). *dismo : Species distribution modeling*. R package version 0.9-3. Cité pages [58](#), [72](#), [101](#) et [125](#).
- HIJMANS, R. J. et van ET TEN, J. (2012). raster : Geographic analysis and modeling with raster data. *R package version*, 1:9–92. Cité page [199](#).
- HINTON, M. A. C. (1926). *Monograph of the Voles & Lemmings (Microtinae) : Living and Extinct*. British Museum (Natural History), London. Cité page [48](#).
- HIRZEL, A., HAUSSE, J., CHESSEL, D. et PERRIN, N. (2002). Ecological-niche factor analysis : how to compute habitat-suitability maps without absence data? *Ecology*, 83(7):2027–2036. Cité page [55](#).
- HOLLAND, J. H. (1975). *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press. Cité page [57](#).
- HOLT, R. D. (1985). Population dynamics in two-patch environments : some anomalous consequences of an optimal habitat distribution. *Theoretical Population Biology*, 28(2):181–208. Cité page [10](#).

- HOLZKNECHT, S. (1989). The Markham languages of Papua New Guinea. A history of the Austronesian languages of the Markham and Ramu valleys, Papua New Guinea. Canberra : Research School of Pacific Studies. *The Australian National University. (Pacific Linguistics Series C-115)*. Cité page [144](#).
- HOLZMANN, I., AGOSTINI, I., DEMATTEO, K., ARETA, J. I., MERINO, M. L. et DI BITETTI, M. S. (2015). Using species distribution modeling to assess factors that determine the distribution of two parapatric howlers (*Alouatta* spp.) in South America. *International Journal of Primatology*, 36(1):18–32. Cité page [70](#).
- HOPFIELD, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558. Cité page [55](#).
- HOSMER, D. W. J. et LEMESHOW, S. (1989). *Assessing the fit of the model*. In : *Applied Logistic Regression*. Wiley, New York. Cité page [55](#).
- HOULDER, D., HUTCHINSON, M., NIX, H. et MC MAHON, J. (2000). ANUCLIM (Version 5.1). *Centre for Resource and Environmental Studies, Australian National University, Canberra*. Cité page [55](#).
- HUBBELL, S. P. (2006). Neutral theory and the evolution of ecological equivalence. *Ecology*, 87(6):1387–1398. Cité page [10](#).
- HUTCHINSON, G. (1957). Cold Spring Harbor Symposium on Quantitative Biology. *Concluding Remarks*, 22:415–427. Cité pages [10](#), [15](#), [19](#), [25](#), [29](#) et [125](#).
- JACKSON, S. T. et OVERPECK, J. T. (2000). Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, 26(sp4):194–220. Cité page [18](#).
- JAYNES, E. T. (1957). Information Theory and Statistical Mechanics. *Physical Review*, 106:620–630. Cité page [55](#).
- JEUNESSE, C. (1997). *Pratiques funéraires au néolithique ancien : sépultures et nécropoles des sociétés danubiennes (5500-4900 av. J.-C.)*. Editions Errance. Cité page [250](#).
- JEUNESSE, C. (2004). Les traditions funéraires du Néolithique moyen en Europe centrale dans le cadre du système funéraire danubien. *Cahiers de l'Association pour la promotion de la recherche archéologique en Alsace*, 20:3–26. Cité pages [228](#) et [250](#).
- JIANG, Y., METZ, C. E. et NISHIKAWA, R. M. (1996). A receiver operating characteristic partial area index for highly sensitive diagnostic tests. *Radiology*, 201(3):745–750. Cité page [78](#).

- JOHNSON, C., JOHNSON, L., RICHARDS, C. et BEASLEY, V. (2002). Predicting the occurrence of amphibians : an assessment of multiple-scale models. *Predicting species occurrences : issues of accuracy and scale*, pp. 157–170. Cité pages 27 et 68.
- JOHNSON, J. B. et OMLAND, K. S. (2004). Model selection in ecology and evolution. *Trends in Ecology & Evolution*, 19(2):101–108. Cité page 110.
- JONES, P., BRIFFA, K., OSBORN, T., LOUGH, J., VAN OMMEN, T., VINTHER, B., LUTERBACHER, J., WAHL, E., ZWIERS, F., MANN, M. *et al.* (2009). High-resolution palaeoclimatology of the last millennium : a review of current status and future prospects. *The Holocene*, 19(1):3–49. Cité page 188.
- JONES, P. D. et MANN, M. E. (2004). Climate over past millennia. *Reviews of Geophysics*, 42(2). Cité page 188.
- JOST, A., LUNT, D., KAGEYAMA, M., ABE-OUCHI, A., PEYRON, O., VALDES, P. et RAMSTEIN, G. (2005). High-resolution simulations of the last glacial maximum climate over Europe : a solution to discrepancies with continental palaeoclimatic reconstructions? *Climate Dynamics*, 24(6):577–590. Cité page 90.
- JUNGCLAUS, J. H., LORENZ, S. J., TIMMRECK, C., REICK, C. H., BROVKIN, V., SIX, K., SEGSCHEIDER, J., GIORGETTA, M. A., CROWLEY, T. J., PONGRATZ, J., KRIVOVA, N. A., VIEIRA, L. E., SOLANKI, S. K., KLOCKE, D., BOTZET, M., ESCH, M., GAYLER, V., HAAK, H., RADDATZ, T. J., ROECKNER, E., SCHNUR, R., WIDMANN, H., CLAUSSEN, M., STEVENS, B. et MAROTZKE, J. (2010). Climate and carbon-cycle variability over the last millennium. *Climate of the Past Discussions*, 6(3):1009–1044. Cité pages 199, 201 et 202.
- JÜPTNER, H. (1983). *Tropenmedizinische Untersuchung der Eipo im zentralen Bergland von Irian Jaya (West-Neuguinea), Indonesien*, volume 12. D. Reimer. Cité page 145.
- KAGEYAMA, M., BRACONNOT, P., BOPP, L., CAUBEL, A., FOIJOLS, M.-A., GUILYARDI, E., KHO-DRI, M., LLOYD, J., LOMBARD, F., MARIOTTI, V. *et al.* (2013). Mid-Holocene and Last Glacial Maximum climate simulations with the IPSL model—part I : comparing IPSL_CM5A to IPSL_CM4. *Climate Dynamics*, 40(9-10):2447–2468. Cité page 65.
- KAGEYAMA, M., NEBOUT, N. C., SEPULCHRE, P., PEYRON, O., KRINNER, G., RAMSTEIN, G. et CAZET, J.-P. (2005). The Last Glacial Maximum and Heinrich Event 1 in terms of climate and vegetation around the Alboran Sea : a preliminary model-data comparison. *Comptes Rendus Geoscience*, 337(10):983–992. Cité page 65.
- KANG, S. M., HELD, I. M., FRIERSON, D. M. et ZHAO, M. (2008). The response of the ITCZ to extratropical thermal forcing : Idealized slab-ocean experiments with a GCM. *Journal of Climate*, 21(14):3521–3532. Cité page 65.

- KARLSSON, G. (2000). *Iceland's 1100 years : the history of a marginal society*. C. Hurst & Co. Publishers. Cité page [192](#).
- KARR, J. R. (1981). *Rationale and techniques for samplify avian habitats : introduction*. IN : *the use of multivariate estatistics in studies of wildlife habitat*. Thèse de doctorat, ed D : E : Capen. US : Department of Agriculture, Forest Service, General Technical Report RM'87. Rocky Mountain Forest and Range Experiment Station. Fort Collins CO. Cité page [14](#).
- KAYSER, M., BRAUER, S., CORDAUX, R., CASTO, A., LAO, O., ZHIVOTOVSKY, L. A., MOYSE-FAURIE, C., RUTLEDGE, R. B., SCHIEFENHOEVEL, W., GIL, D. *et al.* (2006). Melanesian and Asian origins of Polynesians : mtDNA and Y chromosome gradients across the Pacific. *Molecular Biology and Evolution*, 23(11):2234–2244. Cité pages [143](#) et [144](#).
- KAYSER, M., BRAUER, S., WEISS, G., SCHIEFENHÖVEL, W., UNDERHILL, P., SHEN, P., OEFNER, P., TOMMASEO-PONZETTA, M. et STONEKING, M. (2003). Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. *The American Journal of Human Genetics*, 72(2):281–302. Cité pages [144](#) et [145](#).
- KAYSER, M., BRAUER, S., WEISS, G., SCHIEFENHÖVEL, W., UNDERHILL, P. A. et STONEKING, M. (2001). Independent histories of human Y chromosomes from Melanesia and Australia. *The American Journal of Human Genetics*, 68(1):173–190. Cité page [142](#).
- KAYSER, M., BRAUER, S., WEISS, G., UNDERHILL, P., ROEWER, L., SCHIEFENHÖVEL, W. et STONEKING, M. (2000). Melanesian origin of Polynesian Y chromosomes. *Current Biology*, 10(20):1237–1246. Cité page [144](#).
- KAYSER, M., CHOI, Y., van OVEN, M., MONA, S., BRAUER, S., TRENT, R. J., SUARKIA, D., SCHIEFENHÖVEL, W. et STONEKING, M. (2008). The impact of the Austronesian expansion : evidence from mtDNA and Y chromosome diversity in the Admiralty Islands of Melanesia. *Molecular Biology and Evolution*, 25(7):1362–1374. Cité page [144](#).
- KEARNEY, M. (2006). Habitat, environment and niche : what are we modelling? *Oikos*, 115(1):186–191. Cité pages [25](#), [28](#) et [30](#).
- KELLER, C. (1991). Vikings in the West Atlantic : a model of Norse Greenlandic medieval society. *Acta Archaeologica*, 61(1990):126–141. Cité page [194](#).
- KELLY, R. L. (1983). Hunter-gatherer mobility strategies. *Journal of Anthropological Research*, pp. 277–306. Cité page [124](#).
- KELLY, R. L. (1992). Mobility/sedentism : concepts, archaeological measures, and effects. *Annual Review of Anthropology*, pp. 43–66. Cité page [124](#).

- KERTÉSZ, R. et SÜMEGI, P. (2001). Theories, critiques and a model : why did the expansion of the Körös-Starčevo culture stop in the centre of the Carpathian Basin. *From the Mesolithic to the Neolithic*, pp. 225–46. Cité page [227](#).
- KIMURA, M. *et al.* (1968). Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626. Cité page [10](#).
- KIRCH, P. V. (1997). *The Lapita peoples : Ancestors of the Oceanic world. The Peoples of South-East Asia and the Pacific*. Blackwell, Cambridge, MA. Cité page [143](#).
- KIRKMAN, H., KIDSON, C. et KENNEDY, M. (1968). Variants of human glucose-6-phosphate dehydrogenase. Studies of samples from New Guinea. *Hereditary disorders of erythrocyte metabolism. Grune and Stratton, New York*, pp. 126–146. Cité page [145](#).
- KLINE, M. A. et BOYD, R. (2010). Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society of London B : Biological Sciences*, 277(1693):2559–2564. Cité page [68](#).
- KOCH, L. (1945). *The east Greenland ice*. Meddelelser om Grønland, 130 (3), 373. Cité page [193](#).
- KOTTEK, M., GRIESER, J., BECK, C., RUDOLF, B. et RUBEL, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3):259–263. Cité page [136](#).
- KOWALSKI, K. (1995). Lemmings (Mammalia, Rodentia) as indicators of temperature and humidity in the European Quaternary. *Acta Zoologica Cracoviensia*, 38(1):85–94. Cité page [48](#).
- KREUZ, A. (2007). Archaeobotanical perspectives on the beginning of agriculture north of the Alps. *The origins and spread of domestic plants in southwest Asia and Europe. Left Coast Press, Walnut Creek*, pp. 259–294. Cité pages [227](#) et [231](#).
- KREUZ, A., MARINOVA, E., SCHÄFER, E. et WIETHOLD, J. (2005). A comparison of early Neolithic crop and weed assemblages from the Linearbandkeramik and the Bulgarian Neolithic cultures : differences and similarities. *Vegetation History and Archaeobotany*, 14(4):237–258. Cité pages [226](#) et [227](#).
- KROGH, K. . (1982). *Erik den Rødes Grønland*. København : Nationalmuseets Forlag. Cité pages [88](#) et [196](#).
- KUIJPERS, A., ABRAHAMSEN, N., HOFFMANN, G., HÜHNERBACH, V., KONRADI, P., KUNZENDORF, H., MIKKELSEN, N., THIEDE, J. et WEINREBE, W. (1999). Climate change and the Viking-age fjord environment of the Eastern Settlement, South Greenland. *Geology of*

- Greenland Survey Bulletin-Review of Greenland Activities*, 183:61–67. Cité pages [195](#), [196](#) et [222](#).
- KUIJPERS, A. et MIKKELSEN, N. (2009). Geological records of changes in wind regime over south Greenland since the Medieval Warm Period : a tentative reconstruction. *Polar Record*, 45(1):1–8. Cité page [196](#).
- LARSON, G., ALBARELLA, U., DOBNEY, K., ROWLEY-CONWY, P., SCHIBLER, J., TRESSET, A., VIGNE, J.-D., EDWARDS, C. J., SCHLUMBAUM, A., DINU, A. *et al.* (2007). Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences*, 104(39):15276–15281. Cité page [226](#).
- LASSEN, S. J., KUIJPERS, A., KUNZENDORF, H., HOFFMANN-WIECK, G., MIKKELSEN, N. et KONRADI, P. (2004). Late-Holocene Atlantic bottom-water variability in Igaliku Fjord, South Greenland, reconstructed from foraminifera faunas. *The Holocene*, 14(2):165–171. Cité page [196](#).
- LE HOUÉROU, H. N. *et al.* (1989). *The grazing land ecosystems of the African Sahel*. Springer-Verlag. Cité pages [174](#) et [175](#).
- LEATHWICK, J., ELITH, J., FRANCIS, M., HASTIE, T. et TAYLOR, P. (2006). Variation in demersal fish species richness in the oceans surrounding New Zealand : an analysis using boosted regression trees. *Marine Ecology Progress Series*, 321:267–281. Cité page [55](#).
- LEATHWICK, J., MOILANEN, A., FRANCIS, M., ELITH, J., TAYLOR, P., JULIAN, K., HASTIE, T. et DUFFY, C. (2008). Novel methods for the design and evaluation of marine protected areas in offshore waters. *Conservation Letters*, 1(2):91–102. Cité page [55](#).
- LEHMANN, A., MCC OVERTON, J. et LEATHWICK, J. R. (2002). GRASP : generalized regression analysis and spatial prediction. *Ecological Modelling*, 157(2-3):189–207. Cité page [55](#).
- LEK, S., DELACOSTE, M., BARAN, P., DIMOPOULOS, I., LAUGA, J. et phane AULAGNIER, S. (1996). Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling*, 90(1):39–52. Cité page [55](#).
- LEVAVASSEUR, G. (2012). *Modélisation environnementale sous contraintes climatiques : approches statistiques pour le pergélisol et la végétation*. Thèse de doctorat, Versailles-St Quentin en Yvelines. Cité page [92](#).
- LEVAVASSEUR, G., VRAC, M., ROCHE, D. M., PAILLARD, D., MARTIN, A. et VANDENBERGHE, J. (2011). Present and LGM permafrost from climate simulations : contribution of statistical downscaling. *Climate of the Past*, 7(4):1225–1246. Cité page [66](#).
- LEVINS, R. (1968). *Evolution in changing environments : some theoretical explorations*. Numéro 2. Princeton University Press. Cité page [125](#).

- LEVINS, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the ESA*, 15(3):237–240. Cité page [11](#).
- LI, Z. et CONIL, S. (2003). Transient response of an atmospheric GCM to North Atlantic SST anomalies. *Journal of Climate*, 16(23):3993–3998. Cité page [230](#).
- LIM, H. C., RAHMAN, M. A., LIM, S. L. H., MOYLE, R. G. et SHELDON, F. H. (2011). Revisiting Wallaces Haunt : Coalescent simulations and comparative niche modeling reveal historical mechanisms that promoted avian population divergence in the Malay archipelago. *Evolution*, 65(2):321–334. Cité pages [42](#), [67](#) et [91](#).
- LIU, C., BERRY, P. M., DAWSON, T. P. et PEARSON, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3):385–393. Cité pages [33](#) et [72](#).
- LOBO, J. M., JIMÉNEZ-VALVERDE, A. et REAL, R. (2008). AUC : a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17(2):145–151. Cité page [78](#).
- LOISELLE, B. A., HOWELL, C. A., GRAHAM, C. H., GOERCK, J. M., BROOKS, T., SMITH, K. G. et WILLIAMS, P. H. (2003). Avoiding Pitfalls of Using Species Distribution Models in Conservation Planning. *Conservation Biology*, 17(6):1591–1600. Cité page [41](#).
- LOSOS, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology letters*, 11(10):995–1003. Cité page [34](#).
- LUBBOCK, J. (1865). *Pre-historic times*. Williams and Norgate. Cité page [226](#).
- LYNCH, J., ROSS, M. et CROWLEY, T. (2002). *The oceanic languages*, volume 1. Psychology Press. Cité pages [143](#) et [147](#).
- LYNNERUP, N. (1998). The Greenland Norse : a biological-anthropological study. *Meddelelser om Grønland - Man & Society*, 24:1–149. Cité page [193](#).
- MABOGUNJE, A. L. et RICHARDS, P. (1979). The land and peoples of West Africa. *History of West Africa*, 1:1–32. Cité page [132](#).
- MACARTHUR, R. H. (1958). Population Ecology of Some Warblers of Northeastern Coniferous Forests. *Ecology*, 39(4):599–619. Cité pages [19](#), [20](#) et [297](#).
- MACARTHUR, R. H. (1967). *The theory of island biogeography*, volume 1. Princeton University Press. Cité page [10](#).
- MACARTHUR, R. H. et WILSON, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution*, pp. 373–387. Cité page [10](#).

- MACE, R. et HOLDEN, C. J. (2005). A phylogenetic approach to cultural evolution. *Trends in Ecology & Evolution*, 20(3):116 – 121. Cité page [134](#).
- MACE, R. et PAGEL, M. (1995). A latitudinal gradient in the density of human languages in North America. *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 261(1360):117–121. Cité pages [37](#), [133](#) et [146](#).
- MAILLEFER, J.-M. (2005). Les Vikings en Russie. *Mémoires/Histoire*, (1):106–129. Cité page [189](#).
- MANEL, S., WILLIAMS, H. C. et ORMEROD, S. J. (1999). Alternative methods for predicting species distribution : an illustration with Himalayan river birds. *Journal of Applied Ecology*, 36(5):734–747. Cité pages [55](#) et [72](#).
- MANEN, C., MARCHAND, G. et CARVALHO, A. (2007). Le Néolithique ancien de la péninsule Ibérique : vers une nouvelle évaluation du mirage africain ? In *Le Néolithique ancien de la péninsule Ibérique : vers une nouvelle évaluation du mirage africain ?*, pp. 133–151. Société préhistorique française. Cité page [232](#).
- MANN, M. E. (2002). Medieval climatic optimum. *Encyclopedia of Global Environmental Change*, 1:514–516. Cité pages [188](#), [202](#) et [299](#).
- MANN, M. E., ZHANG, Z., RUTHERFORD, S., BRADLEY, R. S., HUGHES, M. K., SHINDELL, D., AMMANN, C., FALUVEGI, G. et NI, F. (2009). Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. *Science*, 326(5957):1256–1260. Cité page [202](#).
- MARCHAND, G. (2003). Les zones de contact Mésolithique/Néolithique dans l’ouest de la France : définition et implications. *Muita gente, poucas antas ? Origens, espaços e contextos do Megalitismo : Actas do II Coloquio Internacional sobre Megalitismo*, pp. 181–197. Cité page [226](#).
- MARCHAND, G. et MANEN, C. (2006). Le rôle du Néolithique ancien méditerranéen dans la néolithisation de l’Europe atlantique. In ADRAHP-PSO, éditeur : *6e Rencontres Méridionales de Préhistoire Récente, 2004*, pp. 213–232. Périgueux. Cité pages [226](#), [236](#) et [241](#).
- MARMION, M., PARVAINEN, M., LUOTO, M., HEIKKINEN, R. K. et THUILLER, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, 15(1):59–69. Cité pages [110](#), [112](#) et [115](#).
- MARSKE, K. A., LESCHEN, R. A. et BUCKLEY, T. R. (2011). Reconciling phylogeography and ecological niche models for New Zealand beetles : Looking beyond glacial refugia. *Molecular Phylogenetics and Evolution*, 59(1):89–102. Cité page [42](#).

- MARTI, O., BRACONNOT, P., BELLIER, J., BENSHILA, R., BONY, S., BROCKMANN, P., CADULE, P., CAUBEL, A., DENVIL, S., DUFRESNE, J. *et al.* (2005). The new IPSL climate system model : IPSL-CM4. *Note du Pôle de Modélisation, IPSL*, 26:1–86. Cité page 202.
- MARTÍNEZ-MEYER, E., PETERSON, A. T., SERVÍN, J. I. et KIFF, L. F. (2006). Ecological niche modelling and prioritizing areas for species reintroductions. *Oryx*, 40(04):411–418. Cité page 52.
- MASSA, C. (2010). L'implantation médiévale scandinave de la côte sud-ouest du Groenland comme la limite du modèle agro-pastoral importé d'Europe du nord : Implications paléoenvironnementales. *Sciences Humaines Combinées, LISIT - revue électronique de l'ED 491*(n°5, Limite / Limites). Cité pages 192 et 193.
- MASSA, C. (2012). *Variabilité climatique holocène et impacts anthropiques historiques en zone subarctique : Étude multiparamètre de la séquence sédimentaire du lac d'Igaliku (Groenland)*. Thèse de doctorat, Université de Franche-Comté. Cité pages 193, 195 et 197.
- MASSA, C., BICHET, V., GAUTHIER, É., PERREN, B. B., MATHIEU, O., PETIT, C., MONNA, F., GIRAUDEAU, J., LOSNO, R. et RICHARD, H. (2012). A 2500 year record of natural and anthropogenic soil erosion in South Greenland. *Quaternary Science Reviews*, 32:119–130. Cité page 196.
- MASTRORILLO, S., LEK, S., DAUBA, F. et BELAUD, A. (1997). The use of artificial neural networks to predict the presence of small-bodied fish in a river. *Freshwater Biology*, 38(2):237–246. Cité page 55.
- MATEO, R. G., FELICÍSIMO, Á. M. et MUNOZ, J. (2010). Effects of the number of presences on reliability and stability of MARS species distribution models : the importance of regional niche variation and ecological heterogeneity. *Journal of Vegetation Science*, 21(5):908–922. Cité page 55.
- MAYR, E. (1942). *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press. Cité page 9.
- MCCULLAGH, P. et NELDER, J. A. (1989). Generalized linear models. In *Monographs on Statistics and Applied Probability*. Cité pages 54 et 55.
- MCCULLOCH, W. et PITTS, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133. Cité pages 55 et 56.
- MCGOVERN, T. H. (1985a). The Arctic Frontier of Norse Greenland. In : *The Archaeology of Frontiers and Boundaries*. *The Archaeology of Frontiers and Boundaries*, pp. 275–323. Cité page 195.

- MCGOVERN, T. H. (1985b). Contributions to the Paleoeconomy of Norse Greenland. *Acta Archaeologica*, 54(1983):73–122. Cité page [194](#).
- MCGOVERN, T. H. (1991). Climate, Correlation, and Causation in Norse Greenland. *Arctic Anthropology*, 28(2):77–100. Cité page [195](#).
- MCGOVERN, T. H., PERDIKARIS, S., EINARSSON, Á. et SIDELL, J. (2006). Coastal connections, local fishing, and sustainable egg harvesting : patterns of Viking Age inland wild resource use in Mývatn district, Northern Iceland. *Environmental Archaeology*, 11:187–205. Cité pages [193](#), [194](#) et [195](#).
- MCGOVERN, T. H., VÉSTEINSSON, O., FRÍÐRIKSSON, A., CHURCH, M., LAWSON, I., SIMPSON, I. A., EINARSSON, A., DUGMORE, A., COOK, G., PERDIKARIS, S., EDWARDS, K. J., THOMSON, A. M., ADDERLEY, W. P., NEWTON, A., LUCAS, G., EDVARSSON, R., ALDREDED, O. et DUNBAR, E. (2007). Landscapes of Settlement in Northern Iceland : Historical Ecology of Human Impact and Climate Fluctuation on the Millennial Scale. *American Anthropologist*, 109(1): 27–51. Cité pages [196](#) et [197](#).
- MCLEMAN, R. et SMIT, B. (2006). Migration as an adaptation to climate change. *Climatic Change*, 76(1-2):31–53. Cité page [3](#).
- MELTON, T., PETERSON, R., REDD, A. J., SAHA, N., SOFRO, A., MARTINSON, J. et STONEKING, M. (1995). Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *American Journal of Human Genetics*, 57(2):403. Cité pages [143](#) et [147](#).
- MERRIWETHER, D. A., FRIEDLAENDER, J. S., MEDIAVILLA, J., MGONE, C., GENTZ, F. et FERRELL, R. E. (1999). Mitochondrial DNA variation is an indicator of Austronesian influence in Island Melanesia. *American Journal of Physical Anthropology*, 110:243–270. Cité page [179](#).
- MIKKELSEN, N., KUIJPERS, A. et ARNEBORG, J. (2008). The Norse in Greenland and late Holocene sea-level change. *Polar Record*, 44:45–50. Cité page [196](#).
- MILLS, W. J. (2003). *Exploring polar frontiers : a historical encyclopedia*, volume 1. ABC-CLIO. Cité page [192](#).
- MOHEN, J.-P. (2006). Climat et néolithisation de l'Europe méditerranéenne. *Comptes Rendus Palevol*, 5(1):453–462. Cité page [226](#).
- MONROE, B. P., NAKAZAWA, Y. J., REYNOLDS, M. G. et CARROLL, D. S. (2014). Estimating the geographic distribution of human Tanapox and potential reservoirs using ecological niche modeling. *International Journal of Health Geographics*, 13(1):1–9. Cité page [70](#).
- MORRISON, M. L. (2001). A proposed research emphasis to overcome the limits of wildlife-habitat relationship studies. *The Journal of Wildlife Management*, pp. 613–623. Cité page [14](#).

- MOUNTAIN, M.-J. (1993). Bones, hunting and predation in the Pleistocene of northern Sahul. *Sahul in Review : Pleistocene Archaeology in Australia, New Guinea and Island Melanesia*, pp. 123–130. Cité page 142.
- MÜLLER, I., BOCKARIE, M., ALPERS, M. et SMITH, T. (2003). The epidemiology of malaria in Papua New Guinea. *Trends in Parasitology*, 19(6):253–259. Cité page 145.
- NETTLE, D. (1996). Language diversity in West Africa : An ecological approach. *Journal of Anthropological Archaeology*, 15(4):403–438. Cité pages 132, 134, 174, 175, 176 et 185.
- NETTLE, D. (1998). Explaining Global Patterns of Language Diversity. *Journal of Anthropological Archaeology*, 17(4):354–374. Cité pages 37, 134, 135, 146, 171, 172, 174, 175 et 298.
- NETTLE, D. (2009). Ecological influences on human behavioural diversity : a review of recent findings. *Trends in Ecology & Evolution*, 24(11):618–624. Cité page 134.
- NEW, M., LISTER, D., HULME, M. et MAKIN, I. (2002). A high-resolution data set of surface climate over global land areas. *Climate Research*, 21(1):1–25. Cité pages 93 et 94.
- NIX, H. A. (1986). A biogeographic analysis of Australian elapid snakes. In : Atlas of Elapid Snakes of Australia. (Ed.) R. Longmore. *Australian Flora and Fauna Series*, 7:4–15. Cité pages 55, 67 et 68.
- NOGUÉS-BRAVO, D. (2009). Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, 18(5):521–531. Cité pages 34 et 42.
- NORRIS, D. (2014). Model thresholds are more important than presence location type : Understanding the distribution of lowland tapir (*Tapirus terrestris*) in a continuous Atlantic forest of southeast Brazil. *Tropical Conservation Science*, 7(3):529–547. Cité pages 69 et 72.
- NORRIS, K. (2004). Managing threatened species : the ecological toolbox, evolutionary theory and declining-population paradigm. *Journal of Applied Ecology*, 41(3):413–426. Cité page 41.
- NORTH, D. C. et THOMAS, R. P. (1977). The first economic revolution. *The Economic History Review*, 30(2):229–241. Cité page 227.
- NOURY, A. et GALIPAUD, J.-C. (2011). *Les Lapita, nomades du Pacifique*. IRD Editions. Cité pages 143 et 144.
- O'CONNELL, J. F. et ALLEN, J. (2004). Dating the colonization of Sahul (Pleistocene Australia-New Guinea) : a review of recent research. *Journal of Archaeological Science*, 31(6):835–853. Cité pages 140 et 142.

- ODUM, Eugene P. & Odum, H. (1959). Fundamentals of ecology, 564 pp. *Philadelphia : WB Saunders Company*. Cité page [24](#).
- OLSON, D. M., DINERSTEIN, E., WIKRAMANAYAKE, E. D., BURGESS, N. D., POWELL, G. V., UNDERWOOD, E. C., D'AMICO, J. A., ITOUA, I., STRAND, H. E., MORRISON, J. C. *et al.* (2001). Terrestrial Ecoregions of the World : A New Map of Life on Earth A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, 51(11):933–938. Cité pages [137](#), [139](#) et [140](#).
- OPPENHEIMER, S. et RICHARDS, M. (2001). Fast trains, slow boats, and the ancestry of the Polynesian islanders. *Science Progress*, 84(3):157–181. Cité pages [143](#) et [144](#).
- ORTEGA-HUERTA, M. A. et PETERSON, A. T. (2004). Modelling spatial patterns of biodiversity for conservation prioritization in North-eastern Mexico. *Diversity and Distributions*, 10(1):39–54. Cité page [41](#).
- OTTE, M. et NOIRET, P. (2001). Le Mésolithique du Bassin Pannonien et la formation du Rubané. *L'Anthropologie*, 105(3):409–419. Cité pages [226](#) et [241](#).
- PAGEL, M. et MACE, R. (2004). The cultural wealth of nations. *Nature*, 428(6980):275–278. Cité page [4](#).
- PASCHOU, P., DRINEAS, P., YANNAKI, E., RAZOU, A., KANAKI, K., TSETOS, E., PADMANABHUNI, S. S., MICHALODIMITRAKIS, M., RENDA, M. C., PAVLOVIC, S. *et al.* (2014). Maritime route of colonization of Europe. *Proceedings of the National Academy of Sciences*, 111(25):9211–9216. Cité page [227](#).
- PATTERSON, W. P., DIETRICH, K. A., HOLMDEN, C. et ANDREWS, J. T. (2010). Two millennia of North Atlantic seasonality and implications for Norse colonies. *Proceedings of the National Academy of Sciences*, 107(12):5306–5310. Cité page [195](#).
- PAWLAK, Z. (1991). *Rough sets : Theoretical aspects of reasoning about data*, volume 9. Springer. Cité page [55](#).
- PAWLEY, A. (2002). The Austronesian dispersal : languages, technologies and people. *Examining the farming/language dispersal hypothesis*. Cambridge : McDonald Institute for Archaeological Research, pp. 251–273. Cité page [140](#).
- PAWLEY, A. (2005). The chequered career of the Trans New Guinea hypothesis : recent research and its implications. *Papuan pasts : Cultural, linguistic and biological histories of Papuan-speaking peoples*, pp. 67–108. Cité page [146](#).
- PEARCE, J. et FERRIER, S. (2000). An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling*, 128(2–3):127–147. Cité page [55](#).

- PEARSON, R. G. et DAWSON, T. P. (2003). Predicting the impacts of climate change on the distribution of species : are bioclimate envelope models useful? *Global Ecology and Biogeography*, 12(5):361–371. Cité page 26.
- PEARSON, R. G. et DAWSON, T. P. (2004). Bioclimate envelope models : what they detect and what they hide—response to Hampe (2004). *Global Ecology and Biogeography*, 13(5):471–473. Cité page 72.
- PEARSON, R. G., RAXWORTHY, C. J., NAKAMURA, M. et PETERSON, A. T. (2007). Predicting species distributions from small numbers of occurrence records : a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, 34(1):102–117. Cité page 72.
- PEREIRA, J. et ITAMI, R. M. (1991). GIS-based habitat modeling using logistic multiple regression- A study of the Mt. Graham red squirrel. *Photogrammetric Engineering and Remote Sensing*, 57(11):1475–1486. Cité page 55.
- PERLÈS, C. (2010). Grèce et Balkans : deux voies de pénétration distinctes du Néolithique en Europe. *La révolution néolithique dans le monde. Paris : CNRS edition*, pp. 263–281. Cité page 227.
- PETERSON, A. T. (2006). Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics*, 3:59–72. Cité pages 24 et 26.
- PETERSON, A. T. (2011). Ecological niche conservatism : A time-structured review of evidence. *Journal of Biogeography*, 38(5):817–827. Cité pages 34 et 41.
- PETERSON, A. T. (2013). Biodiversity Informatics Training Curriculum - Ecological Niche Modeling. Cité page 50.
- PETERSON, A. T., BAUER, J. T. et MILLS, J. N. (2004). Ecologic and geographic distribution of filovirus disease. *Emerging Infectious Diseases*, 10:40–47. Cité page 41.
- PETERSON, A. T. et COHOON, K. (1999). Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling*, 117(1):159–164. Cité pages 13 et 26.
- PETERSON, A. T. et LI, X. (2014). Niche-based projections of wetlands shifts with marine intrusion from sea level rise : an example analysis for North Carolina. *Environmental Earth Sciences*, pp. 1–12. Cité page 41.
- PETERSON, A. T., MOSES, L. M. et BAUSCH, D. G. (2014). Mapping Transmission Risk of Lassa Fever in West Africa : The Importance of Quality Control, Sampling Bias, and Error Weighting. *PloS ONE*, 9(8):e100711. Cité pages 41 et 75.
- PETERSON, A. T. et NAKAZAWA, Y. (2008). Environmental data sets matter in ecological niche modelling : an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, 17(1):135–144. Cité page 71.

- PETERSON, A. T., ORTEGA-HUERTA, M. A., BARTLEY, J., SÁNCHEZ-CORDERO, V., SOBERÓN, J., BUDDEMEIER, R. H. et STOCKWELL, D. R. (2002). Future projections for Mexican faunas under global climate change scenarios. *Nature*, 416(6881):626–629. Cité pages 13 et 55.
- PETERSON, A. T., PAPEŞ, M. et EATON, M. (2007). Transferability and model evaluation in ecological niche modeling : a comparison of GARP and Maxent. *Ecography*, 30(4):550–560. Cité pages 33, 67 et 91.
- PETERSON, A. T., PAPEŞ, M. et SOBERÓN, J. (2008a). Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, 213(1):63–72. Cité pages 56, 57, 78, 79, 110 et 119.
- PETERSON, A. T. et ROBINS, C. R. (2003). Using Ecological-Niche Modeling to Predict Barred Owl Invasions with Implications for Spotted Owl Conservation. *Conservation Biology*, 17(4):1161–1165. Cité page 41.
- PETERSON, A. T., SOBERON, J., PEARSON, R., ANDERSON, R., MARTINEZ-MEYER, E., NAKAMURA, M. et ARAÚJO, M. (2011). *Ecological Niches and Geographic Distributions*. Princeton Univ Pr. Cité pages 18, 30, 31, 32, 47 et 72.
- PETERSON, A. T., SOBERÓN, J. et SÁNCHEZ-CORDERO, V. (1999). Conservatism of Ecological Niches in Evolutionary Time. *Science*, 285(5431):1265–1267. Cité page 34.
- PETERSON, A. T., STEWART, A., MOHAMED, K. I. et ARAÚJO, M. B. (2008b). Shifting global invasive potential of European plants with climate change. *PLoS ONE*, 3(6):e2441. Cité page 52.
- PETERSON, A. T. et VIEGLAIS, D. A. (2001). Predicting species snvasions using ecological niche modeling : new approaches from bioinformatics attack a pressing problem a new approach to ecological niche modeling, based on new tools drawn from biodiversity informatics, is applied to the challenge of predicting potential species' invasions. *BioScience*, 51(5):363–371. Cité page 41.
- PÉTREQUIN, P. et PÉTREQUIN, A.-M. (1993). *Écologie d'un outil : la hache de pierre en Irian Jaya (Indonésie)*, volume 12. CNRS. Cité page 142.
- PHILLIPS, S. J., ANDERSON, R. P. et SCHAPIRE, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3):231–259. Cité page 57.
- PHILLIPS, S. J., DUDÍK, M. et SCHAPIRE, R. E. (2004). A maximum entropy approach to species distribution modeling. In *Proceedings of the twenty-first international conference on Machine learning*, p. 83. ACM. Cité pages 55, 56 et 57.

- PINHASI, R. (2006). Neolithic skull shapes and demic diffusion : a bioarchaeological investigation into the nature of the Neolithic transition. *Documenta Praehistorica*, 33:61–70. Cité page [226](#).
- PIPERNO, D. R. (2011). The origins of plant cultivation and domestication in the New World tropics. *Current Anthropology*, 52(S4):S453–S470. Cité page [142](#).
- POCHEVILLE, A. (2010). *La Niche Ecologique : Concepts, Modèles, Applications*. Thèse de doctorat, Ecole Normale Supérieure de Paris-ENS Paris. Cité page [10](#).
- POULOS, H. M., CHERNOFF, B., FULLER, P. L. et BUTMAN, D. (2012). Ensemble forecasting of potential habitat for three invasive fishes. *Aquatic Invasions*, 7(1):59–72. Cité page [115](#).
- PRENTICE, M. L. et HOPE, G. S. (2007). Climate of Papua. *The Ecology of Papua*, 1:479–494. Cité page [136](#).
- PRICE, D. H. (1989). *Atlas of world cultures : a geographical guide to ethnographic literature*. Sage Publications. Cité page [37](#).
- PRICE, T. D. (2000). *Europe's first farmers*. Cambridge University Press. Cité page [227](#).
- PRINZING, A., DURKA, W., KLOTZ, S. et BRANDL, R. (2002). Geographic variability of ecological niches of plant species : are competition and stress relevant ? *Ecography*, 25(6):721–729. Cité pages [27](#) et [28](#).
- PROST, S., SMIRNOV, N., FEDOROV, V. B., SOMMER, R. S., STILLER, M., NAGEL, D., KNAPP, M. et HOFREITER, M. (2010). Influence of climate warming on arctic mammals ? New insights from ancient DNA studies of the collared lemming *Dicrostonyx torquatus*. *PLoS ONE*, 5(5):e10447. Cité page [48](#).
- PULLIAM, H. (2000). On the relationship between niche and distribution. *Ecology Letters*, 3(4):349–361. Cité pages [24](#), [26](#) et [30](#).
- PULLIAM, H. R. (1988). Sources, Sinks, and Population Regulation. *The American Naturalist*, 132(5):652–661. Cité page [10](#).
- PURUGGANAN, M. D. et FULLER, D. Q. (2009). The nature of selection during plant domestication. *Nature*, 457(7231):843–848. Cité page [142](#).
- QIAO, H., SOBERÓN, J. et PETERSON, A. T. (2015). No Silver Bullets in Correlative Ecological Niche Modeling : Insights from Testing Among Many Potential Algorithms for Niche Estimation. *Methods in Ecology and Evolution*. Cité page [110](#).
- RAES, N. et ter STEEGE, H. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, 30(5):727–736. Cité page [78](#).

- RANDI, E. (2008). Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology*, 17(1):285–293. Cité page [226](#).
- RAPOPORT, E. H. (1982). *Areography : geographical strategies of species*, volume 1. New York : Pergamon Press. Cité page [133](#).
- RAXWORTHY, C. J., INGRAM, C. M., RABIBISOA, N. et PEARSON, R. G. (2007). Applications of ecological niche modeling for species delimitation : a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. *Systematic Biology*, 56(6):907–923. Cité page [55](#).
- REESINK, G., SINGER, R. et DUNN, M. (2009). Explaining the linguistic diversity of Sahul using population models. *PLoS Biology*, 7(11):e1000241. Cité pages [146](#) et [147](#).
- RENFREW, C. (1990). *Archaeology and language : the puzzle of Indo-European origins*. CUP Archive. Cité page [226](#).
- RENFREW, C. (1996). Language families and the spread of farming. *The origins and spread of agriculture and pastoralism in Eurasia*, pp. 70–92. Cité page [226](#).
- RICHERSON, P. et BOYD, R. (2005). Not by Genes Alone : How Culture Transformed Human Evolution. *Univeristy of Chicago Press, Chicago*. Cité page [38](#).
- RICHERSON, P. J., BOYD, R. et BETTINGER, R. L. (2009). Cultural innovations and demographic change. *Human Biology*, 81(3):211–235. Cité page [38](#).
- RIDGEWAY, G. (2006). gbm : Generalized boosted regression models. *R package version*, 1(3). Cité page [55](#).
- RIGAUD, S., D'ERRICO, F. et VANHAEREN, M. (2015). Ornaments Reveal Resistance of North European Cultures to the Spread of Farming. *PLoS ONE*, 10(4):e0121166. Cité pages [241](#) et [243](#).
- RISSLER, L. J. et APODACA, J. J. (2007). Adding more ecology into species delimitation : ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Systematic Biology*, 56(6):924–942. Cité page [69](#).
- RIVOLLAT, M., MENDISCO, F., PEMONGE, M.-H., SAFI, A., SAINT-MARC, D., BRÉMOND, A., COUTURE-VESCHAMBRE, C., ROTTIER, S. et DEGUILLOUX, M.-F. (2015). When the Waves of European Neolithization Met : First Paleogenetic Evidence from Early Farmers in the Southern Paris Basin. *PLoS ONE*, 10(4):e0125521. Cité page [243](#).
- ROBERTS, R. G., JONES, R. et SMITH, M. A. (1990). Thermoluminescence dating of a 50,000-year-old human occupation site in northern Australia. *Nature*, 345:153 – 156. Cité page [140](#).

- ROBERTSON, M. P., CAITHNESS, N. et VILLET, M. H. (2001). A PCA-based modelling technique for predicting environmental suitability for organisms from presence records. *Diversity and Distributions*, 7(1-2):15–27. Cité page 55.
- ROBERTSON, M. P., VILLET, M. H. et PALMER, A. R. (2004). A fuzzy classification technique for predicting species' distributions : applications using invasive alien plants and indigenous insects. *Diversity and Distributions*, 10(5-6):461–474. Cité pages 55 et 72.
- RODE, A. L. S. et LIEBERMAN, B. S. (2005). Using environmental niche modeling to study the Late Devonian biodiversity crisis. *Developments in Palaeontology and Stratigraphy*, 20:93–179. Cité page 42.
- RÖSCH, M., EHLMANN, O., HERRMANN, L., SCHULZ, E., BOGENRIEDER, A., GOLDAMMER, J. P., HALL, M., PAGE, H. et SCHIER, W. (2002). An experimental approach to Neolithic shifting cultivation. *Vegetation History and Archaeobotany*, 11(1):143–154. Cité page 226.
- ROSE, J. P. et TODD, B. D. (2014). Projecting Invasion Risk of Non-Native Watersnakes (*Nerodia fasciata* and *Nerodia sipedon*) in the Western United States. *PloS ONE*, 9(6): e100277. Cité page 41.
- ROSS, M. *et al.* (2005). Pronouns as a preliminary diagnostic for grouping Papuan languages. *Papuan pasts : Cultural, linguistic and biological histories of Papuan-speaking peoples*, pp. 15–66. Cité pages 146, 148 et 177.
- ROSSIGNOL-STRICK, M. (1999). The Holocene climatic optimum and pollen records of sapropel 1 in the eastern Mediterranean, 9000–6000BP. *Quaternary Science Reviews*, 18(4):515–530. Cité page 245.
- ROTTIER, S., MORDANT, C., CHAMBON, P. et THEVENET, C. (2005). Découverte de plus d'une centaine des sépultures du Néolithique moyen à Gurgy, les Noisats (Yonne). *Bulletin de la Société préhistorique française*, pp. 641–645. Cité page 243.
- ROUGET, M., RICHARDSON, D. M., NEL, J. L., LE MAITRE, D. C., EGOH, B. et MGIDI, T. (2004). Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions*, 10(5-6):475–484. Cité page 41.
- ROURA-PASCUAL, N., BROTONS, L., PETERSON, A. T. et THUILLER, W. (2009). Consensual predictions of potential distributional areas for invasive species : a case study of Argentine ants in the Iberian Peninsula. *Biological Invasions*, 11(4):1017–1031. Cité page 109.
- ROYER, A. (2013). *Etude paléoenvironnementale et paléoclimatique du Pléistocène supérieur du Sud-Ouest de la France, à partir d'analyses comparées d'associations fauniques et de biogéochimies effectuées sur les micromammifères*. Thèse de doctorat, EPHE. Cité page 49.

- RUSHTON, S. P., ORMEROD, S. J. et KERBY, G. (2004). New paradigms for modelling species distributions? *Journal of Applied Ecology*, 41(2):193–200. Cité page [41](#).
- RUWENDE, C. et HILL, A. (1998). Glucose-6-phosphate dehydrogenase deficiency and malaria. *Journal of Molecular Medicine*, 76(8):581–588. Cité page [145](#).
- SAMY, A. M., van de SANDE, W. W. J., FAHAL, A. H. et PETERSON, A. T. (2014). Mapping the Potential Risk of Mycetoma Infection in Sudan and South Sudan Using Ecological Niche Modeling. *PLoS Neglected Tropical Diseases*, 8(10):e3250. Cité page [41](#).
- SÁNCHEZ, M. C., ESPEJO, F. J. J., VALLEJO, M. D. S., BAO, J. F. G., CARVALHO, A. F., MARTINEZ-RUIZ, F., GAMIZ, M. R., FLORES, J.-A., PAYTAN, A., SÁEZ, J. A. L. *et al.* (2012). The Mesolithic–Neolithic transition in southern Iberia. *Quaternary Research*, 77(2):221–234. Cité page [232](#).
- SÁNCHEZ-CORDERO, V., CIRELLI, V., MUNGUIAL, M. et SARKAR, S. (2005). Place prioritization for biodiversity content using species ecological niche modeling. *Biodiversity Informatics*, 2. Cité page [41](#).
- SANDEL, B. et DANGREMOND, E. M. (2012). Climate change and the invasion of California by grasses. *Global Change Biology*, 18(1):277–289. Cité pages [41](#) et [51](#).
- SATTLER, T., BONTADINA, F., HIRZEL, A. H. et ARLETTAZ, R. (2007). Ecological niche modelling of two cryptic bat species calls for a reassessment of their conservation status. *Journal of Applied Ecology*, 44(6):1188–1199. Cité page [55](#).
- SCHIEFENHÖVEL (2014). Human Ethological Perspectives on Prehistoric Adaptation and Dispersal in the Central Highlands of New Guinea. In : Sanz, N. (Ed.), Human Evolution, Adaptation and Social Developments. *UNESCO, World Heritage Programs*, 39:235–254. Cité pages [141](#), [142](#), [144](#), [147](#), [148](#), [150](#), [181](#) et [182](#).
- SCHIEFENHÖVEL, W. (1982). Results of ethnomedical fieldwork among the Eipo, Daerah Jayawijaya, hian Jaya, with spocial reference to Eaditional birthgiying. *Medika (Jaküta) I l*. Cité page [145](#).
- SCHIEFENHÖVEL, W. (2001). Kampf, Krieg und Versöhnung bei den Eipo im Bergland von West-Neuguinea–Zur Evolutionsbiologie und Kulturanthropologie aggressiven Verhaltens. In Fikentscher, W. (Eds.) *Begegnung und Konflikt - eine kulturanthropologische Bestandsaufnahme. Bayerische Akademie der Wissenschaften, Philosophisch-Historische Klasse, Abhandlungen, Neue Folge, Heft 120. C.H. Beck, München*, pp. 169–186. Cité page [176](#).
- SCHIEFENHÖVEL, W. (2013). Biodiversity through Domestication. Examples from New Guinea. *Revue d'ethnoécologie*, (3). Cité page [142](#).

- SCHIEFENHOVEL, W. et BLUM, P. (2007). Insects : forgotten and rediscovered as food. Entomophagy among the Eipo, highlands of west New Guinea, and other tribal societies. *Consuming the Inedible : Neglected Dimensions of Food Choice*. Berghahn, New York, p. 16. Cité page [143](#).
- SCHLEDERMANN, P. et MC CULLOUGH, K. (2003). Inuit-Norse Contact in the Smith Sound Region. *Contact, Continuity and Collapse the Norse Colonization of the North Atlantic*. ed. James H. Barrett. Belgium : Brepols, pp. 183–205. Cité page [223](#).
- SCHOENER, T. W. (1968). The Anolis lizards of Bimini : resource partitioning in a complex fauna. *Ecology*, 49(4):704–726. Cité pages [125](#), [176](#) et [178](#).
- SCHÖLKOPF, B., PLATT, J. C., SHAW-TAYLOR, J., SMOLA, A. J. et WILLIAMSON, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7): 1443–1471. Cité page [55](#).
- SCHÖLKOPF, B., SMOLA, A. J., WILLIAMSON, R. C. et BARTLETT, P. L. (2000). New support vector algorithms. *Neural Computation*, 12(5):1207–1245. Cité page [55](#).
- SEPULCHRE, P., GILLES, R., KAGEYAMA, M., MARIAN, V., KRINNER, G., SÁNCHEZ GOÑI, M. F. et D'ERRICO, F. (2007). {H4} abrupt event and late Neanderthal presence in Iberia. *Earth and Planetary Science Letters*, 258(1–2):283–292. Cité page [43](#).
- SEPULCHRE, P., SCHUSTER, M., RAMSTEIN, G., KRINNEZR, G., GIRARD, J.-F., VIGNAUD, P. et BRUNET, M. (2008). Evolution of Lake Chad Basin hydrology during the mid-Holocene : A preliminary approach from lake to climate modelling. *Global and Planetary Change*, 61(1):41–48. Cité page [230](#).
- SERVONNAT, J. (2010). *Variabilité climatique en Atlantique Nord au cours du dernier millénaire : évaluation de l'influence du forçage solaire avec le modèle IPSLCM4*. Thèse de doctorat, Versailles-St Quentin en Yvelines. Cité pages [92](#), [93](#), [188](#), [202](#), [203](#) et [219](#).
- SHINDELL, D. T., SCHMIDT, G. A., MANN, M. E., RIND, D. et WAPLE, A. (2001). Solar forcing of regional climate change during the Maunder Minimum. *Science*, 294(5549):2149–2152. Cité page [202](#).
- SIDDALL, M., ROHLING, E. J., ALMOGI-LABIN, A., HEMLEBEN, C., MEISCHNER, D., SCHMELZER, I. et SMEED, D. (2003). Sea-level fluctuations during the last glacial cycle. *Nature*, 423(6942):853–858. Cité page [64](#).
- SIDERA, I. (1997). Le mobilier en matières dures animales en milieu funéraire Cerny : symbolisme et socio-économie. La culture de Cerny. In CONSTANTIN, C., MORDANT, D. et SIMONIN, D., éditeurs : *Nouvelle économie, nouvelle société au Néolithique, Actes du colloque international de Nemours 1994*, pp. 499–513. Mémoires du musée de préhistoire d'Ile-de-France, Nemours. Cité page [228](#).

- SIDERA, I. (2000). Animaux domestiques, bêtes sauvages et objets en matière animales du Rubané au Michelsberg. *Gallia Préhistoire*, 42:107–194. Cité page [228](#).
- SIDERA, I. (2008). Rubané, Villeneuve-Saint-Germain et Cardial. Filiation des industries osseuses. In CONSTANTIN, C., BURNEZ-LANOTTE, C., ALLARD, P. et ILLET, M., éditeurs : *Fin des traditions danubiennes dans le néolithique du Bassin parisien et de la Belgique (5100-4700 av. J.-C.)*, pp. 209–219. Cité page [241](#).
- SILVA, D. P., GONZALEZ, V. H., MELO, G. A., LUCIA, M., ALVAREZ, L. J. et DE MARCO, P. (2014). Seeking the flowers for the bees : integrating biotic interactions into niche models to assess the distribution of the exotic bee species *Lithurgus huberi* in South America. *Ecological Modelling*, 273:200–209. Cité page [55](#).
- SKOV, F. et SVENNING, J.-C. (2004). Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography*, 27(3):366–380. Cité page [41](#).
- SMITH, B. D. (2011). The cultural context of plant domestication in eastern North America. *Current Anthropology*, 52(S4):71–84. Cité page [142](#).
- SMITH, E. A. (2001). On the coevolution of cultural, linguistic, and biological diversity. *On biocultural diversity : Linking language, knowledge, and the environment*, pp. 95–117. Cité page [133](#).
- SMITHERS, J. et SMIT, B. (1997). Human adaptation to climatic variability and change. *Global Environmental Change*, 7(2):129–146. Cité page [3](#).
- SOBERÓN, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10(12):1115–1123. Cité pages [26](#) et [125](#).
- SOBERÓN, J. et NAKAMURA, M. (2009). Niches and distributional areas : concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, 106(Supplement 2):19644–19650. Cité pages [18](#), [26](#), [27](#), [32](#), [39](#) et [125](#).
- SOBERON, J. et PETERSON, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2(0):1–10. Cité pages [17](#), [24](#) et [26](#).
- SOLHEIM, W. G. (1984). The Nusantara hypothesis : the origin and spread of Austronesian speakers. *Asian Perspectives*, pp. 77–88. Cité page [144](#).
- SPATZ, H. (1997). La nécropole du Néolithique moyen (Hinkelstein, Grossgartach) de Trebur (Gross-Gerau, Hesse). Le Néolithique danubien et ses marges entre Rhin et Seine, Actes du 22e colloque interrégional sur le Néolithique. *Cahiers de l'Association pour la promotion de la recherche archéologique en Alsace*, (suppl.):157–170. Cité page [228](#).

- SPETH, J. D. (1987). Early hominid subsistence strategies in seasonal habitats. *Journal of Archaeological Science*, 14(1):13–29. Cité page [3](#).
- STARK, R. et BAINBRIDGE, W. S. (1987). *A theory of religion*, volume 2. Lang New York. Cité page [3](#).
- STEVENS, G. C. (1989). The latitudinal gradient in geographical range : how so many species coexist in the tropics. *American Naturalist*, pp. 240–256. Cité page [133](#).
- STIGALL, A. L. (2012). Using ecological niche modelling to evaluate niche stability in deep time. *Journal of Biogeography*, 39(4):772–781. Cité page [42](#).
- STOCKWELL, D. (1999a). The GARP modelling system : problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13(2):143–158. Cité pages [55](#), [57](#) et [58](#).
- STOCKWELL, D. R. (1999b). Genetic algorithms II. In *Machine learning methods for ecological applications*, pp. 123–144. Springer. Cité pages [55](#), [56](#) et [57](#).
- STOCKWELL, D. R. et NOBLE, I. (1992). Induction of sets of rules from animal distribution data : a robust and informative method of data analysis. *Mathematics and Computers in Simulation*, 33(5-6):385–390. Cité pages [55](#) et [57](#).
- STOCKWELL, D. R. et PETERSON, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148(1):1–13. Cité pages [60](#), [61](#) et [110](#).
- STONEKING, M., JORDE, L. B., BHATIA, K. et WILSON, A. C. (1990). Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics*, 124(3):717–733. Cité page [145](#).
- STRAHLER, A. H. (1980). The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment*, 10(2):135–163. Cité page [55](#).
- STRINGER, C. (2003). Human evolution : out of Ethiopia. *Nature*, 423(6941):692–695. Cité page [14](#).
- STUIVER, M., GROOTES, P. M. et BRAZIUNAS, T. F. (1995). The GISP2 $\delta^{18}\text{O}$ climate record of the past 16,500 years and the role of the sun, ocean, and volcanoes. *Quaternary Research*, 44(3):341–354. Cité page [195](#).
- SUMMERHAYES, G. R., FORD, A., DENNELL, R. et PORR, M. (2014). Late Pleistocene Colonisation and Adaptation in New Guinea Implications for Modelling Modern Human Behaviour. *Southern Asia, Australia and the Search for Human Origins*, p. 213. Cité pages [140](#) et [142](#).

- SVEINSSON, E. Ó., ÞÓRÐARSON, M. et HALLDÓRSSON, Ó. (1935). *Eyrbyggja saga. Brands þátrr örva. Eiríks saga rauða. Grœnlendinga saga. Grœnlendinga þátrr*. Hið íslenska fornritafélag. Cité page 189.
- SVENNING, J.-C., FLØJGAARD, C., MARSKE, K. A., NOGUÉS-BRAVO, D. et NORMAND, S. (2011). Applications of species distribution modeling to paleobiology. *Quaternary Science Reviews*, 30(21-22):2930–2947. Cité pages 42 et 51.
- SVENNING, J.-C. et SKOV, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters*, 7(7):565–573. Cité pages 24 et 26.
- SWADLING, P. (1983). *How long have people been in the Ok Tedi Impact Region ?* Numéro 8. PNG national museum. Cité page 146.
- SWADLING, P. (1997). Changing shorelines and cultural orientations in the Sepik-Ramu, Papua New Guinea : Implications for Pacific prehistory. *World Archaeology*, 29(1):1–14. Cité pages 146 et 147.
- SWINGEDOUW, D., ORTEGA, P., MIGNOT, J., GUILYARDI, E., MASSON-DELMOTTE, V., BUTLER, P. G., KHODRI, M. et SÉFÉRIAN, R. (2015). Bidecadal North Atlantic ocean circulation variability controlled by timing of volcanic eruptions. *Nature Communications*, 6:6545. Cité pages 202 et 203.
- TEAM, Q. D. (2009). *QGIS Geographic Information System*. Open Source Geospatial Foundation. Cité page 86.
- TEAM, R. C. (2013). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Cité page 165.
- TERRILL, A. (2003). Linguistic stratigraphy in the central Solomon Islands : Lexical evidence of early Papuan/Austronesian interaction. *The Journal of the Polynesian Society*, pp. 369–401. Cité page 179.
- THUILLER, W. (2003). BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, 9(10):1353–1362. Cité pages 54, 55 et 115.
- THUILLER, W., ARAÚJO, M. B. et LAVOREL, S. (2003). Generalized models vs. classification tree analysis : predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, 14(5):669–680. Cité page 55.
- THUILLER, W., MIDGLEY, G., ROUGETI, M. et COWLING, R. (2006). Predicting patterns of plant species richness in megadiverse South Africa. *Ecography*, 29(5):733–744. Cité page 55.

- THUILLER, W., GEORGES, D. et ENGLER, R. (2013). biomod2 : Ensemble platform for species distribution modeling. *R package version*, 2(7):r560. Cité pages 58, 109 et 115.
- THUILLER, W., LAFOURCADE, B., ENGLER, R. et ARAÚJO, M. B. (2009). BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography*, 32(3):369–373. Cité pages 58 et 114.
- THUILLER, W., LAVOREL, S. et ARAÚJO, M. (2005a). Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology and Biogeography*, 14(4):347–357. Cité page 41.
- THUILLER, W., RICHARDSON, D. M., PYSEK, P., MIDGLEY, G. F., HUGHES, G. O. et ROUGET, M. (2005b). Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, 11(12):2234–2250. Cité page 41.
- TIERNEY, J., BIELER, L. *et al.* (1967). *Diculi liber de mensura orbis terrae*, volume 6. Dublin Institute for Advances Studies. Cité page 192.
- TORRENCE, R., NEALL, V., DOELMAN, T., RHODES, E., MCKEE, C., DAVIES, H., BONETTI, R., GUGLIELMETTI, A., MANZONI, A., ODDONE, M. *et al.* (2004). Pleistocene colonisation of the Bismarck Archipelago : new evidence from West New Britain. *Archaeology in Oceania*, 39(3):101–130. Cité page 142.
- TRAVIS, J., BROOKER, R. et DYTHAM, C. (2005). The interplay of positive and negative species interactions across an environmental gradient : insights from an individual-based simulation model. *Biology Letters*, 1(1):5–8. Cité page 45.
- VÁCLAVÍK, T. et MEENTEMEYER, R. K. (2009). Invasive species distribution modeling (iSDM) : Are absence data and dispersal constraints needed to predict actual distributions ? *Ecological Modelling*, 220(23):3248–3258. Cité page 55.
- Van der VAART, A. W. (1998). *Asymptotic statistics*. Cambridge University Press, U.K. Cité pages 125 et 178.
- VAN VALEN, L. (1977). The red queen. *American Naturalist*, pp. 809–810. Cité page 33.
- VANDERWAL, J., FALCONI, L., JANUCHOWSKI, S., SHOO, L. et STORLIE, C. (2011). SDMTTools : Species distribution modelling tools : Tools for processing data associated with species distribution modelling exercises. *R package version*, 1(5). Cité page 125.
- VAPNIK, V. (1995). *The nature of statistical learning theory*. Springer-Verlag, New York. Cité page 55.
- VAUGHAN, I. P. et ORMEROD, S. J. (2003). Improving the Quality of Distribution Models for Conservation by Addressing Shortcomings in the Field Collection of Training Data. *Conservation Biology*, 17(6):1601–1611. Cité page 69.

- VAYSSIÈRES, M. P., PLANT, R. E. et ALLEN-DIAZ, B. H. (2000). Classification trees : An alternative non-parametric approach for predicting species distributions. *Journal of Vegetation Science*, 11(5):679–694. Cité page 55.
- VELICHKOVSKY, B. M. et RUMBAUGH, D. M. (2013). *Communicating meaning : The evolution and development of language*. Psychology Press. Cité page 146.
- VERHART, L. B. (1990). Stone age bone and antler points as indicators for “Social Territories” in the European mesolithic. *Contributions to the Mesolithic in Europe*, pp. 139–151. Cité page 34.
- VIGNE, J.-D. (1993). Domestication ou appropriation pour la chasse : histoire d’un choix socio-culturel depuis le Néolithique. L’exemple des cerfs (*Cervus*). *Exploitation des animaux sauvages à travers le temps*, pp. 201–220. Cité page 226.
- VIGNE, J.-D. (1998). Faciès culturels et sous-système technique de l’acquisition des ressources animales. Application au Néolithique ancien méditerranéen. *Rencontres méridionales de Préhistoire récente, 1996, Arles*, pp. 27–45. Cité page 226.
- VIGNE, J.-D. (2000). Les débuts néolithiques de l’élevage des ongulés au Proche-Orient et en Méditerranée : acquis récents et questions. In Guilaine J. (ed.), *Premiers paysans du monde. Naissances des agricultures*, pp. 143–168. Séminaire du Collège de France, éditions Errance, Paris. Cité page 226.
- VIGNE, J.-D. (2008). Zooarchaeological aspects of the Neolithic diet transition in the Near East and Europe, and their putative relationships with the Neolithic demographic transition. In *The Neolithic demographic transition and its consequences*, pp. 179–205. Springer. Cité page 226.
- VOLTERRA, V. (1928). Variations and fluctuations of the number of individuals in animal species living together. *J. Cons. Int. Explor. Mer*, 3(1):3–51. Cité page 19.
- VRAC, M., MARBAIX, P., PAILLARD, D., NAVEAU, P. et al. (2007). Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe. *Climate of the Past*, 3(4):669–682. Cité pages 66 et 92.
- WALKER, P. A. (1990). Modelling wildlife distributions using a geographic information system : kangaroos in relation to climate. *Journal of Biogeography*, pp. 279–289. Cité page 55.
- WALLACE, B. L. (2000). *The viking settlement at L’Anse aux meadows*. In : Fitzhugh, W.W., and E. I. Ward (editors). *Vikings : the North Atlantic saga*. Washington : Smithsonian Institution Press : 208–216. Cité page 191.

- WANG, X., HUANG, X., JIANG, L. et QIAO, G. (2010). Predicting potential distribution of chestnut phylloxerid (Hemiptera : Phylloxeridae) based on GARP and Maxent ecological niche models. *Journal of Applied Entomology*, 134(1):45–54. Cité page [110](#).
- WARD, D. F. (2007). Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions*, 9(6):723–735. Cité page [55](#).
- WARREN, D. L., GLOR, R. E. et TURELLI, M. (2008). Environmental niche equivalency versus conservatism : quantitative approaches to niche evolution. *Evolution*, 62(11):2868–2883. Cité pages [34](#) et [125](#).
- WARREN, D. L., GLOR, R. E. et TURELLI, M. (2010). ENMTools : a toolbox for comparative studies of environmental niche models. *Ecography*, 33(3):607–611. Cité pages [125](#) et [231](#).
- WELLER, O., PÉTREQUIN, P., PÉTREQUIN, A.-M. et COUTURAUD, A. (1996). Du sel pour les échanges sociaux.[L'exploitation des sources salées en Nouvelle-Guinée (Irian Jaya, Indonésie)]. *Journal de la Société des Océanistes*, 102(1):3–30. Cité page [150](#).
- WHITE, J. P., CROOK, K. A. et RUXTON, B. P. (1970). Kosipe : a late Pleistocene site in the Papuan highlands. In *Proceedings of the Prehistoric Society*, volume 36, pp. 152–170. Cité page [142](#).
- WHITTAKER, R. H., LEVIN, S. A. et ROOT, R. B. (1973). Niche, habitat, and ecotope. *American Naturalist*, pp. 321–338. Cité pages [10](#) et [15](#).
- WHITTLE, A. W. (1996). *Europe in the Neolithic : the creation of new worlds*. Cambridge University Press. Cité page [227](#).
- WIENS, J. J. (2004). Speciation and ecology revisited : phylogenetic niche conservatism and the origin of species. *Evolution*, 58(1):193–197. Cité pages [13](#), [34](#) et [133](#).
- WIENS, J. J., ACKERLY, D. D., ALLEN, A. P., ANACKER, B. L., BUCKLEY, L. B., CORNELL, H. V., DAMSCHEN, E. I., JONATHAN DAVIES, T., GRYTNES, J.-A., HARRISON, S. P. *et al.* (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, 13(10):1310–1324. Cité page [34](#).
- WIENS, J. J. et DONOGHUE, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, 19(12):639–644. Cité page [133](#).
- WIENS, J. J. et GRAHAM, C. H. (2005). Niche conservatism : integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, pp. 519–539. Cité pages [13](#) et [37](#).
- WILSON, M. F., O'CONNELL, B., BROWN, C., GUINAN, J. C. et GREHAN, A. J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30(1-2):3–35. Cité page [64](#).

- WINROTH, A. (2014). The Age of the Vikings. *Foreign Rights Guide*, p. 17. Cité page [189](#).
- WURM, S. A. (1983). Linguistic prehistory in the New Guinea area. *Journal of Human Evolution*, 12(1):25–35. Cité page [146](#).
- WURM, S. A., VOORHOEVE, C. L. et MCELHANON, K. A. (1975). The trans-New Guinea phylum in general. *New Guinea area languages and language study*, 1:299–322. Cité page [146](#).
- YAMAMURA, N., HIGASHI, M., BEHERA, N. et WAKANO, J. Y. (2004). Evolution of mutualism through spatial effects. *Journal of Theoretical Biology*, 226(4):421–428. Cité page [45](#).
- YEE, T. W. et MITCHELL, N. D. (1991). Generalized additive models in plant ecology. *Journal of Vegetation science*, 2(5):587–602. Cité page [55](#).
- YIOU, P. et MASSON-DELMOTTE, V. (2005). Trends in sub-annual climate variability since the Little Ice Age in western Europe. *Comptes Rendus Geoscience*, 337(10):1001–1012. Cité page [188](#).
- ZAPATA, L., PEÑA-CHOCARRO, L., PÉREZ-JORDÁ, G. et STIKA, H.-P. (2004). Early neolithic agriculture in the Iberian Peninsula. *Journal of World Prehistory*, 18(4):283–325. Cité pages [226](#) et [232](#).
- ZEDER, M. A. (2005). *A view from the Zagros : new perspectives on livestock domestication in the Fertile Crescent*. Cité page [226](#).
- ZEDER, M. A. (2006). Central questions in the domestication of plants and animals. *Evolutionary Anthropology : Issues, News, and Reviews*, 15(3):105–117. Cité page [226](#).
- ZEDER, M. A. (2008). Domestication and early agriculture in the Mediterranean Basin : Origins, diffusion, and impact. *Proceedings of the National Academy of Sciences*, 105(33): 11597–11604. Cité page [226](#).
- ZHANG, L., LIU, S., SUN, P., WANG, T., WANG, G., ZHANG, X. et WANG, L. (2015). Consensus Forecasting of Species Distributions : The Effects of Niche Model Performance and Niche Properties. *PloS ONE*, 10(3). Cité page [115](#).
- ZIEGLER, M., SIMON, M. H., HALL, I. R., BARKER, S., STRINGER, C. et ZAHN, R. (2013). Development of Middle Stone Age innovation linked to rapid climate change. *Nature Communications*, 4:1905. Cité page [38](#).
- ZILHÃO, J. (2014). Early prehistoric navigation in the Western Mediterranean : Implications for the Neolithic transition in Iberia and the Maghreb. *Eurasian Prehistory*, 1-2(11):185–200. Cité page [232](#).
- ZINGG, P. (1990). Akustische artidentifikation von fledermäusen (Mammalia : Chiroptera) in der Schweiz. *Revue suisse de Zoologie*, 97(2):263–294. Cité page [55](#).

- ZVELEBIL, M. (2000). The social context of the agricultural transition in Europe. *Archaeogenetics : DNA and the population prehistory of Europe*, pp. 57–79. Cité page [226](#).
- ZWEIG, M. H. et CAMPBELL, G. (1993). Receiver-operating characteristic (ROC) plots : a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39(4):561–577. Cité page [78](#).



Table des figures

1.1	Estimation de la densité d'une espèce par Krigeage	12
1.2	Distribution géographique d'une espèce à différentes échelles	12
1.3	Deux hypothèses de distributions spatio-temporelles d'espèces du genre <i>Homo</i> basées sur des occurrences	14
1.4	Représentation théorique d'une niche écologique sur le gradient environnemental	16
1.5	Représentation d'une niche écologique à une, deux et trois dimensions environ- nementales	17
1.6	Le diagramme BAM	17
1.7	Différentes niches au sein d'un même habitat, l'exemple des parulines (<i>Seto- phaga sp</i>) de MacArthur [1958]	20
1.8	Représentations théoriques de niches écologiques d'une espèce seule ou en communauté.	20
1.9	Dynamiques évolutives possibles de deux espèces	22
1.10	Impact des interactions biotiques dans les occurrences	23
1.11	Différentes conceptions des ensembles écologiques.	25
1.12	Modélisation de niche par approche corrélative	29
1.13	Positions des niches modélisées dans le diagramme BAM	30
2.1	Pertinence des données de présence par rapport aux données d'absence.	50
2.2	Fonctionnement général des algorithmes prédictifs.	51
2.3	Projections de niches écologiques.	53
2.4	Effet de la taille de l'échantillon sur les prédictions de niches écologiques.	61
2.5	Pertinence des données de présence ou d'absence dans un concept archéolo- gique ou paléontologique.	63

2.6	Matrice de confusion et représentation spatiale schématique des différents états d'un point dans un modèle	74
2.7	La courbe ROC et l'AUC	79
2.8	Représentation graphique du partial ROC ratio	80
3.1	Origine des données d'occurrences archéologiques du peuplement scandinave du Groenland	88
4.1	Méthode de descente d'échelle statistique GAM	96
4.2	Courbes de régressions ayant servi à la descente d'échelle des températures minimales du mois de janvier	99
4.3	Courbes de régressions ayant servi à la descente d'échelle des températures maximales du mois de janvier	100
4.4	Courbes de régressions ayant servi à la descente d'échelle des précipitations	101
4.5	Visualisation des résidus lors de la validation du modèle GAM.	102
4.6	Comparaison de simulations obtenues par régression GAM d'une simulation CRU dégradée avec la simulation CRU d'origine	103
6.1	Limite de la comparaison de pertinence des prédictions	111
6.2	Construction d'un consensus simple	113
6.3	Consensus simple et consensus pondéré	116
6.4	Création d'un consensus élitiste pondéré	119
6.5	Caractère stochastique des algorithmes prédictifs et consensus	120
8.1	Relation entre densités linguistiques et MGS, d'après [Nettle, 1998]	135
8.2	L'archipel de Nouvelle-Guinée	136
8.3	Climats de Nouvelle-Guinée et diagrammes ombrothermiques	138
8.4	Carte des éco-régions de Nouvelle-Guinée	139
8.5	Carte des biomes de Nouvelle-Guinée	140
8.6	Océanie au DMG	141
8.7	Diversité linguistique de Nouvelle-Guinée	147
8.8	Occurrences utilisées pour l'ELNM en Nouvelle-Guinée	149
8.9	Variations de pertinence des algorithmes prédictifs pour les 525 modélisations de niches écolinguistiques	152
8.10	Distributions géographiques des niches écolinguistiques papoues	161
8.11	Distributions géographiques des niches écolinguistiques austronesiennes	164
8.12	Analyse en composante principale des groupes linguistiques papous et austro-nésiens de Nouvelle-Guinée	166
8.13	Comparaison de l'aire linguistique et de la niche écolinguistique de populations théoriques.	168
8.14	Potentiel écolinguistique de Nouvelle-Guinée	173
8.15	Durée de croissance optimale des végétaux de Nouvelle-Guinée	175

8.16 Arbres phyloécologiques des groupes linguistiques de Nouvelle-Guinée	177
8.17 Comparaison des positions écologiques occupées par les Papous et par les Austronésiens.	180
8.18 Comparaison de la répartition des langues papoues et austronésiennes avec le potentiel écolinguistique	182
9.1 Anomalies des températures du dernier millénaire, d'après Mann [2002]	188
9.2 Extraits de la tapisserie de Bayeux	190
9.3 Emprise des Scandinaves dans les régions arctiques, subarctiques, et dans le Nord de l'Europe	193
9.4 Localisations des occurrences utilisées pour l'ECNM des Vikings d'Islande, du Groenland et de Terre-Neuve	198
9.5 Prédiction des niches écoculturelles des Scandinaves.	201
9.6 Courbes de variations des températures moyennes annuelles au Groenland selon deux simulations climatiques	203
9.7 Prédiction consensuelle de la niche écoculturelle du 1 ^{er} peuplement viking	207
9.8 Prédiction consensuelle de la niche écoculturelle du 2 nd peuplement viking	208
9.9 Projection consensuelle de la niche écoculturelle du 1 ^{er} peuplement viking dans l'environnement du PAG	209
9.10 Projection consensuelle de la niche écoculturelle du 2 nd peuplement viking dans l'environnement de l'OCM	210
9.11 Analyse en composantes principales des ECNM du Groenland.	211
9.12 Cercles de corrélation des analyses en composantes principales des projections et prédictions des niches écoculturelles des peuplements vikings	212
9.13 Importance des variables environnementales	213
9.14 Histogrammes des variables bioclimatiques et topographiques du 1 ^{er} peuplement	214
9.15 Histogrammes des variables bioclimatiques et topographiques du 2 nd peuplement	215
9.16 Variations entre niches prédites et projections	217
9.17 Histogrammes des variables bioclimatiques et topographiques des 1 ^{er} et 2 nd peuplements	218
9.18 Variations locales des températures et précipitations du XI ^e au XV ^e siècle	219
9.19 Dynamique environnementale séculaire des niches des 1 ^{er} et 2 nd peuplements	221
10.1 Cultures archéologiques majeures identifiées au Néolithique ancien en Europe	227
10.2 Occurrences des sites néolithiques	229
10.3 Cartes des M utilisés pour les ECNM du Néolithique	230
10.4 Prédiction des niches écoculturelles des cultures du Néolithique ancien.	236
10.5 Analyse en composante principale des ECNM des cultures du Néolithique	238
10.6 Histogrammes des variables environnementales des prédictions des niches de l'Impressa et du Cardial	240

10.7	Histogrammes des variables environnementales des prédictions de niches méditerranéennes et des danubiennes	242
10.8	Superposition des niches danubiennes et méditerranéennes	244



Liste des tableaux

1.1	Niches modélisées en fonction de la représentativité de la niche réalisée parmi les occurrences échantillonnées.	33
2.1	Exemples de types de niches pouvant être modélisés par les algorithmes prédictifs.	51
2.2	Prédictions et projections de deux niches dans deux environnements	52
2.3	Liste non exhaustive des techniques utilisées pour la prédiction de niches écologiques	55
2.4	Liste de logiciels dédiés à l'application des algorithmes prédictifs	58
2.5	Liste de variables topographiques	64
2.6	Liste des variables bioclimatiques	67
2.7	Méthodes d'établissement du seuil 1-E.	72
3.1	Avantages et désavantages de l'affiliation et de la collecte d'occurrences automatique	87
4.1	Liste des données environnementales fournies par <i>WorldClim</i>	91
4.2	Variables de la simulation CRU (<i>Climate Research Unit</i>).	93
4.3	Liste des variables de la simulation SGI	94
4.4	Validation statistique de la modélisation GAM	98
4.5	Prédicteurs et formules d'interpolation GAM des précipitations et des températures extrêmes.	98
8.1	Regroupements des ELN selon leurs formes et leur position écologique	170
8.2	Permissivités et potentiels écolinguistiques.	172
8.3	<i>Overlaps</i> entre ELN de Nouvelle-Guinée	178

9.1	Causes de la disparition des Vikings du Groenland	196
9.2	Sélection des algorithmes utilisés pour la reconstruction des niches écoculturelles des Vikings	204
10.1	Sélection et pondération des algorithmes utilisés pour la reconstruction des niches écoculturelles du Néolithique ancien	234
10.2	Nombre de prédictions sélectionnées pour l'ACP de l'ECNM du Néolithique . .	237

Application d'algorithmes prédictifs à l'identification de niches écoculturelles des populations du passé : approche ethnoarchéologique

La géographie des groupes humains résulte d'événements historiques culturels et environnementaux. Notre démarche consiste à identifier des relations cultures/environnements dans des populations actuelles ou historiques bien documentées pour ensuite déceler l'éventuelle présence de phénomènes similaires dans des populations anciennes dont seuls les vestiges archéologiques sont connus. Après avoir passé en revue différents concepts permettant de décrire l'espace écologique occupé par une espèce (ou une population déterminée par un trait spécifique), nous déduisons que le concept de niche est idéal pour mesurer les facteurs environnementaux qui peuvent influencer l'établissement d'une culture en un lieu à une période donnée. Afin d'apprécier les distributions géographiques potentielles de cultures du présent et du passé, nous utilisons la modélisation de niches écoculturelles (ECNM). L'ECNM utilise des algorithmes prédictifs ainsi que des données d'occurrences et environnementales afin d'examiner les possibles influences des facteurs environnementaux dans les trajectoires évolutives des cultures. Nous présentons des résultats issus d'une optimisation de l'ECNM qui consiste à obtenir des données environnementales à très hautes résolutions spatiale et temporelle puis à combiner des prédictions de niche en tenant compte des performances des différents algorithmes prédictifs utilisés. La validité de la méthode que nous proposons est assurée par la fiabilité des occurrences que nous utilisons dans nos référentiels actualiste et historique. Enfin l'analyse statistique de plusieurs niches contemporaines, ou se succédant dans plusieurs phases climatiques, nous permet de les positionner dans l'espace écologique et de discuter de diversité culturelle, de risque écologique, de compétition, de dynamique évolutive et de peuplement.

Mots clefs : *Évolution ; Cultures ; Écologie ; Hommes/Environnements ; Changements climatiques.*

Predictive algorithms applied to the identification of eco-cultural niches of past populations: an ethnoarchaeological approach

The geographic distribution of human populations is the result of both historical contingency and environmental factors. This study identifies culture-environment relations for well-documented present-day and historic populations in order to evaluate whether the same phenomena operated in prehistoric contexts, which are only known from archaeological sites. After reviewing the different concepts used to describe the ecological space occupied by a species (or specific population), it is shown that the niche concept is well-suited for identifying and measuring environmental factors that can influence the distribution of a culture at a particular place and time. In order to better understand the potential distributions of present and past cultures, this study employs the method known as eco-cultural niche modeling (ECNM). ECNM uses predictive algorithms along with occurrence and environmental data in order to examine the possible influences of environmental factors on cultural trajectories. The results presented here are derived from an optimized ECNM approach that permits one to obtain high-resolution environmental data, and that also combines niche predictions by taking into account the performance of the various employed predictive algorithms. The effectiveness of this approach is ensured by the use of reliable occurrence data for both the present-day and historic case studies. Finally, statistical evaluations of multiple contemporaneous niches, as well as successive ones across multiple climatic phases, allow them to be placed in ecological space and examined with respect to cultural diversity, ecological risk, competition, and evolutionary and population dynamics.

Keywords: *Evolution; Cultures; Ecology; Human-Environment Relationship; Climate changes.*
